Kleine Enzyklopädie



Atom- und



Vom Elementarteilchen über Atom, Molekül und Festkörper bis zur Materie im Kosmos wird die ganze unbelebte Materie behandelt.

Kernphysik

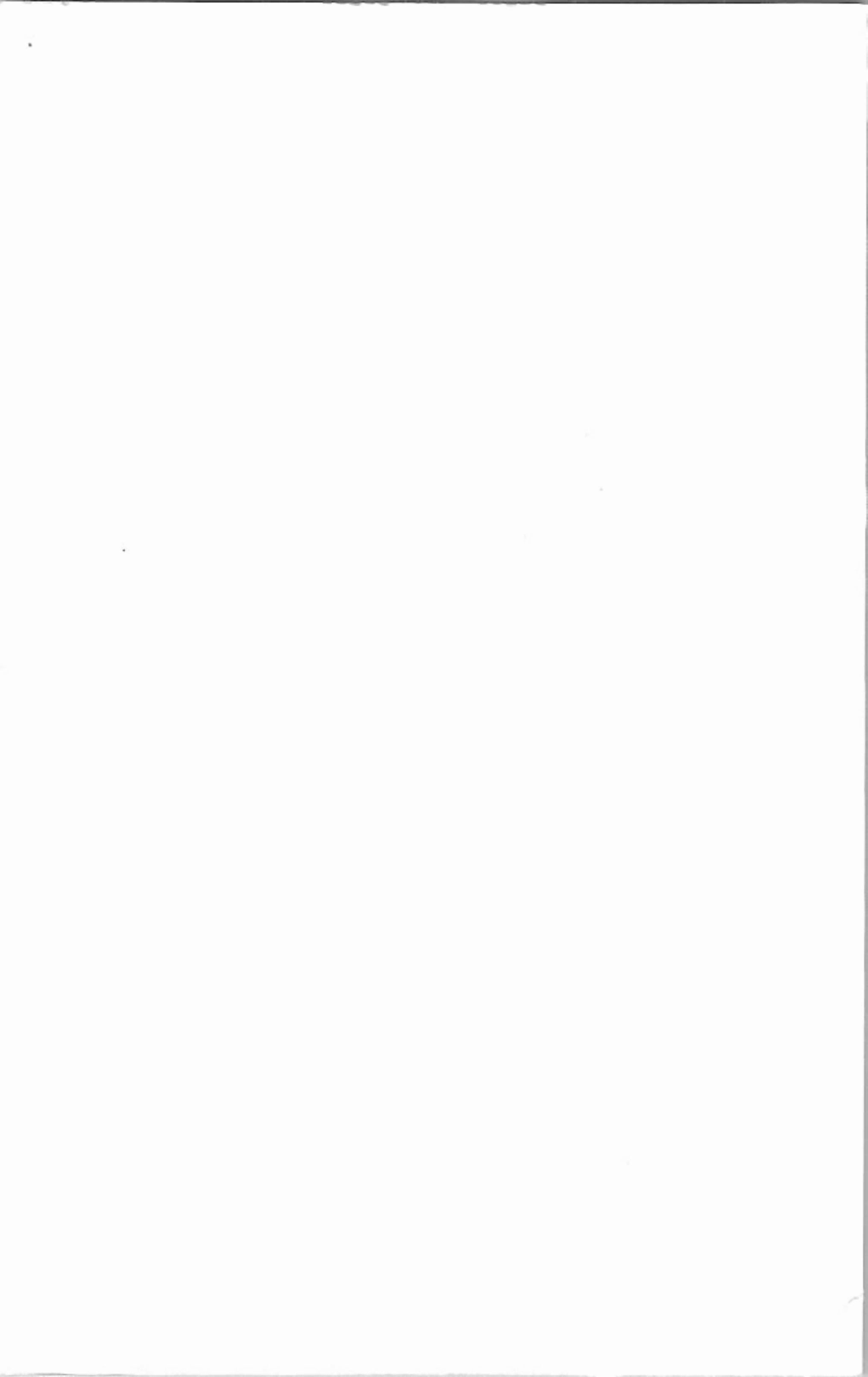
VERLAG HARRI DEUTSCH

Kleine Enzyklopädie
Atom- und Kernphysik

Kleine Enzyklopädie

Atom- und Kernphysik





Kleine Enzyklopädie Atom- und Kernphysik

Kleine Enzyklopädie

Atom- und Kernphysik

Struktur der Materie

über 775 Textabbildungen
und Tabellen sowie 80 Tafeln,
davon 16 mehrfarbig

1983
Verlag Harri Deutsch
Thun · Frankfurt/M.

Herausgeber

Prof. Dr. habil. Christian Weißmantel, Karl-Marx-Stadt (Federführung)
Prof. Dr. habil. Richard Lenk, Karl-Marx-Stadt
Prof. Dr. habil. Wolfgang Forker, Dresden
Doz. Dr. sc. nat. Dietmar Linke, Berlin

Gutachter

Prof. Dr. phil. Rudolf Dobrozemsky, Wien
Prof. Dr. Ing. habil. Paul Görlich, Jena

CIP-Kurztitelaufnahme der Deutschen Bibliothek

Kleine Enzyklopädie. – Thun; Frankfurt/M.:
Deutsch
→ Atom- und Kernphysik
Atom- u. Kernphysik / [Hrsg. Christian
Weißmantel ...]. – Thun; Frankfurt/M.:
Deutsch, 1983
(Kleine Enzyklopädie)
ISBN 3-87144-667-X
NE: Weißmantel, Christian [Hrsg.]

ISBN 3-87144-667-X
Printed in the German Democratic Republic

1. Auflage 1983
Lizenzausgabe für den Verlag Harri Deutsch, Thun

Vorwort

Mit dem vorliegenden Band der Reihe „Kleine Enzyklopädien" wird der Versuch unternommen, die naturwissenschaftlichen Erkenntnisse vom Aufbau und den Eigenschaften der nichtbelebten Materie in einem Band zusammenfassend und überschauend darzustellen. Als verbindender Leitgedanke zieht sich das Bild vom atomistischen Aufbau der Stoffe und von der Quantelung der Kraftfelder durch das Werk hindurch; dementsprechend ist ein einführendes Kapitel den Besonderheiten der atomaren Betrachtungsweise gewidmet. Es schließt sich eine relativ ausführliche Darstellung von Grundlagen der klassischen Physik sowie der Quantentheorie an, auf die der Leser zurückgreifen sollte, wenn sich beim Lesen der späteren Teile prinzipielle Fragen ergeben. Darüberhinaus ist ein systematisches Studium dieser Grunderkenntnisse der Physik demjenigen zu empfehlen, der sich für die Fundamente interessiert, auf denen die physikalische und chemische Beschreibung der materiellen Welt aufbaut.

In den Kapiteln 3. bis 10., die den Hauptinhalt des Werkes bilden, wird die Struktur der physikalischen Materie, beginnend mit den Elementarteilchen und Atomkernen über die Atome und Moleküle bis zu den makroskopischen Stoffzuständen, behandelt. Abgerundet wird die Darstellung durch einen Ausflug in die Kosmophysik, der mit den extragalaktischen Sternsystemen als größten uns bekannten Erscheinungsformen der Materie abschließt.

Ein Studium der Erkenntnisse von der Struktur der Materie ist für den schöpferisch denkenden Menschen in zweifacher Hinsicht bedeutungsvoll: Es festigt sein materialistisches Weltbild und vermittelt ihm zugleich das Rüstzeug für eigene Beiträge bei der Meisterung des wissenschaftlich-technischen Fortschritts. Wenn gegenwärtig die Probleme der Materialökonomie und der Energiewirtschaft als Kernfragen herausgehoben werden, von deren Lösung das weitere Wachstum der Volkswirtschaft bestimmt wird, so muß auch klar sein, daß die entscheidenden Impulse für neue Wege nur aus tiefergehenden Einsichten in die Struktur der Materie gewonnen werden können. Die Halbleitertechnik und Mikroelektronik resultieren aus den Ergebnissen der atomistischen Beschreibung von Festkörpern, und die Bemühungen um die zunehmende Nutzung der Kernenergie wären ohne die wegbereitende Erkundungsforschung auf dem Gebiet der Atomkernphysik undenkbar. Folgerichtig haben sich die Autoren bemüht, den Zusammenhang zwischen der Grundlagenforschung und der konsequenten Nutzung ihrer Ergebnisse herauszuarbeiten.

Hinsichtlich des Niveaus ließen wir uns davon leiten, daß angesichts des wachsenden Bildungsstandes unserer Menschen nicht auf mathematische Hilfsmittel verzichtet werden sollte. Schließlich möchte der interessierte Leser, gleich ob Student oder Facharbeiter, Lehrer oder Wissenschaftler, beim Nachschlagen aus aktuellem Anlaß mehr über die Zusammenhänge erfahren. Auch beim systematischen Studium des Buches, das besonders der interessierten Jugend zu empfehlen ist, sollte die Einheit von Theorie und Praxis beachtet werden. Im übrigen haben sich die Autoren bemüht, den wesentlichen Sachverhalt in Worten so zu formulieren, daß schwierigere mathematische Passagen übersprungen werden können.

Ebenso wie bei den anderen Titeln der Buchreihe sollen eine detaillierte Dezimalklassifikation sowie das sehr umfangreiche Sachregister dem Leser helfen, ein Stichwort einzuordnen bzw. aufzufinden. Zahlreiche Querverweise sowie Formelzähler zielen darauf ab, Zusammenhänge – manchmal auch zwischen anscheinend entfernten Teilgebieten – herzustellen. Schließlich ist die durchgängige Ausstattung mit einer Vielzahl von Textabbildungen, die durch einen Bildtafelteil ergänzt wird, als wesentliches Hilfsmittel zur anschaulichen Untermauerung komplizierter Zusammenhänge sowie zur Vertiefung der Theorie-Praxis-Beziehungen angelegt.

Nahezu täglich wird in den Kommunikationsmitteln von neuen Beobachtungen, Entdekkungen und Anwendungen berichtet, die mit dem Gebiet dieses Buches zusammenhängen. Zweifellos wird der stetig fließende Strom neuer Erkenntnisse manche Überraschung bringen, jedoch ist zu erwarten, daß sich auch die künftigen Ergebnisse der Forschungen zur Struktur der Materie in das gefestigte Bild vom atomaren Aufbau einfügen werden.

Allen Personen und Institutionen, die zur Herausgabe des Werkes beigetragen haben, möchten wir auf diesem Wege danken. Hinweise unserer Leser, die dazu beitragen, das Werk bei eventuellen künftigen Auflagen zu verbessern, werden uns stets willkommen sein.

Frühjahr 1981 Herausgeber und Verlag

Autoren

Dr. rer. nat. Johann Dorschner 10.2.
Prof. Dr. rer. nat. habil. Wolfgang Forker 6.1.
Prof. Dr. sc. nat. Lothar Franke 7.3.4.
Dipl.-Phys. Hans-Erich Fröhlich 10.3.
Dr. rer. nat. habil. Hans-Ulrich Gersch 3.1., 3.4.2., 3.4.4.
Dr. rer. nat. Helmut Giegengack Tafelkonzipierung
Prof. Dr. rer. nat. habil. Gottfried Glöckner 5.4., 6.5.
Dr. sc. nat. Peter Hallpap 7.1., 7.2.2., 7.2.3., 7.3.1.
Prof. Dr. sc. nat. Claus Hamann 8.4. bis 8.7.
Prof. Dr. rer. nat. habil. Ernst Hegenbarth 6.3.6.
Dr. sc. nat. Karl Hehl 4.3.1.
Dr. sc. nat. Jürgen Kleinschmidt 4.1., 4.3.2., 4.3.3.
Prof. Dr.-Ing. habil. Benjamin Kozik 3.5.
Dr. rer. nat. Hubert Langbein 5.1., 5.2., 7.2.1., 7.2.4., 7.2.5.
Prof. Dr. rer. nat. habil. Richard Lenk 1.1.6., 1.2.5., 2., 3.3., 3.4.1., 4.2.
Dr. sc. nat. Dietmar Linke 6.4.2. bis 6.4.6., 7.3,2., 7.3.3.
Dr. rer. nat. Alexander Mende 4.4.
Prof. Dr. rer. nat. habil. Klaus Meyer 6.4.1., 6.4.7., 6.4.8.
Dr. sc. nat. Peter Morzeck 3.4.3.
Dr. sc. nat. Hans Oleak 10.4.
Dr. rer. nat. Günter Reiße 1.1.3., 9.2.
Dr. rer. nat. Heinrich Solbrig 1.1.1., 1.1.2., 1.1.4., 1.1.5., 1.2.1. bis 1.2.4.
Prof. Dr. rer. nat. habil. Eberhard Steger 5.3.
Prof. Dr. rer. nat. habil. Heinz Stiller 9.3.
Dr. rer. nat. Siegfried Ullrich 9.1.
Dr. sc. nat. Heiner Vollstädt 9.3.
Dr. rer. nat. habil. Christian-Ulrich Wagner 10.1.
Dr. rer. nat. Wolfgang Wagner 6.2., 6.3., 6.6.
Prof. Dr. rer. nat. habil. Christian Weißmantel 8.1. bis 8.3., 8.8., 9.4.
Dr. sc. nat. Eberhard Wieczorek 3.2.
Dr. rer. nat. Dieter Wiegel 5.4.7.
Prof. Dr. sc. nat. Bernd Wilhelmi 4.1., 4.3.2., 4.3.3.
Prof. Dr. sc. nat. habil. Helmut Zimmermann 10.3.

Inhalt

1. Einführung

1.1. Das Atom und seine Bestandteile .. 9
1.2. Welle-Teilchen-Dualismus und Plancksches Wirkungsquantum 37

2. Grundlegende physikalische Begriffe und Gesetze

2.1. Bewegung von Teilchen 44
2.2. Elektrische und magnetische Felder 62
2.3. Wellen 71
2.4. Wahrscheinlichkeit und statistische Physik 80
2.5. Grundideen und Aussagen der Quantentheorie 93

3. Elementarteilchen und Atomkerne

3.1. Teilchenbeschleuniger 111
3.2. Elementarteilchen 115
3.3. Eigenschaften und Struktur der Atomkerne 134
3.4. Kernumwandlungen 147
3.5. Kernenergetik.................... 174

4. Elektronen, Lichtquanten und Atome

4.1. Lichtquanten 181
4.2. Struktur und Energiespektren der Atome 196
4.3. Anregung von Atomen und Atomspektren........................ 215
4.4. Spinresonanz 226

5. Moleküle

5.1. Chemische Bindung............... 235
5.2. Struktur der Moleküle 247
5.3. Angeregte Moleküle und Molekülspektren........................ 256
5.4. Makromoleküle 284

6. Zustandsformen der Stoffe

6.1. Makroskopische Teilchensysteme .. 302
6.2. Gase............................ 317
6.3. Flüssigkeiten..................... 327
6.4. Fester Aggregatzustand 352
6.5. Zustandsformen makromolekularer Stoffe 391
6.6. Grenzflächen und disperse Systeme 398

7. Stoffumwandlung

7.1. Grundbegriffe 419
7.2. Kinetik chemischer Reaktionen 425
7.3. Reaktionsverhalten der Stoffe 438

8. Festkörper

8.1. Gitterdynamik der Festkörper...... 471
8.2. Bändermodell der Kristallelektronen 488
8.3. Halbleiter 505
8.4. Leiter.......................... 534
8.5. Supraleiter 548
8.6. Dielektrische und ferroelektrische Erscheinungen in Festkörpern 560
8.7. Magnetische Erscheinungen in Festkörpern 573
8.8. Optische Eigenschaften der Festkörper 587

9. Materie unter extremen Bedingungen

9.1. Plasmazustand 604
9.2. Erzeugung und Messung von Vakuum 629
9.3. Hohe Drücke 632
9.4. Gesteuerte Kernfusion 646

10. Kosmische Erscheinungsformen der Materie

10.1. Planet Erde und erdnaher Raum ... 652
10.2. Erscheinungsformen der Materie im Sonnensystem 668
10.3. Erscheinungsformen der Materie im Milchstraßensystem 686
10.4. Erscheinungsformen der Materie in extragalaktischen Sternsystemen ... 715

Register............................ 734

Quellen

Die Zeichnungen wurden vorwiegend nach Angaben der Autoren angefertigt von:

Jens Borleis, Leipzig · Willi Weitzmann, Leipzig · Joachim Zindler, Leipzig

Wissenschaftliche Aufnahmen und Reproduktionen stellten zur Verfügung:

ADN-Zentralbild, Berlin · Akademie der Wissenschaften der DDR, Zentralinstitut für Isotopen- und Strahlenforschung, Leipzig; Zentralinstitut für Kernforschung, Rossendorf · Akademie-Verlag, Berlin, Redaktion „Wissenschaft und Fortschritt" · Madame Jean Becquerel, Fontainebleau, Frankreich · BI-Archiv, Leipzig · California Institut of Technology Bookstore, Pasadena · CERN, Genf · Cristaltec, Forschungszentrum des CNRS, Grenoble · Deutsche Fotothek, Dresden · VEB Deutscher Verlag für Gundstoffindustrie, Leipzig · Deutsches Museum, München · Dr. Kurt Drescher, Dresden · Dr. Ulrike Ebersbach, Karl-Marx-Stadt · Forschungsinstitut „Manfred von Ardenne", Dresden · Dr. rer. nat. Helmut Giegengack, Karl-Marx-Stadt · Dr. Haberland, Leipzig · Hasselblad, Schweden · Karl-Schwarzschild-Observatorium, Tautenburg · Dipl.-Phys. Ludwig Küchler, Karl-Marx-Stadt · Joachim Müller, Karl-Marx-Stadt · Museum für Naturkunde, Berlin · H.-J. Müssig, Dresden · Dipl.-Phys. Dietrich Pätzold, Berlin · VEB RFT Meßelektronik „Otto Schön", Dresden · Lothar Richter, Jena · Siemens-AG, München · Wolfgang Schröter, Markkleeberg · Prof. Dr. rer. nat. habil. Eberhard Steger, Dresden · Brigitte Stöcker, Potsdam · VEB Verlag Technik, Berlin · Urania-Verlag Leipzig, Jena, Berlin · USIS, Bad Godesberg · Dr. sc. Heiner Vollstädt, Seddin · Friedrich Weimer, Dresden · Prof. Dr. rer. nat. habil. Christian Weißmantel, Karl-Marx-Stadt · Dr. Dieter Wrobel, Sangerhausen · VEB Carl Zeiss, Jena

Hinweise des Verlages zur Lizenzausgabe

Dieses Buch ist in der DDR entwickelt worden und enthält daher neben DDR-spezifischen Beispielen auch einige Wertungen, denen sich nicht jeder Leser anschließen mag. Wir überlassen diese Stellen dem kritischen Urteil des Lesers, zumal unseres Erachtens die naturwissenschaftlichen Aussagen davon nicht berührt werden.

VERLAG HARRI DEUTSCH

1. Einführung

1.1. Das Atom und seine Bestandteile

Die *Atomhypothese* wurde erstmals im 5. Jahrhundert v. u. Z. von einigen Naturphilosophen des antiken Griechenlands (insbesondere von DEMOKRIT) ausgesprochen. Doch erst in der zweiten Hälfte des 19. Jahrhunderts setzte sich die Auffassung allgemein durch, daß die Atome aufgrund ihrer unterschiedlichen Eigenschaften die Vielfalt an chemischen und physikalischen Verhaltensweisen der Stoffe hervorbringen. Chemische Eigenschaften, Druck und Wärme fanden so ihre atomistischen Begründungen. Dennoch blieb es den ersten Jahrzehnten des 20. Jahrhunderts vorbehalten, die Frage nach den Ursachen der unterschiedlichen Eigenschaften der Atome, also die Frage nach der Struktur des Atoms, zu beantworten. Im Zusammenhang damit wurde die *Quantenphysik* entwickelt. Sie ist die Gesamtheit der in atomaren Dimensionen gültigen physikalischen Gesetzmäßigkeiten.

1.1.1. Entstehung eines experimentell begründeten Atomkonzepts

Kennzeichnend für diesen Teil der Wissenschaftsentwicklung ist der Übergang von einer spekulativen materialistischen Naturphilosophie zur Naturwissenschaft. Die Nutzung des Experiments zur Überprüfung des Wahrheitsgehalts von Hypothesen über den Bau der Natur war der entscheidende Faktor in diesem Prozeß.

1.1.1.1. Antike Atomistik und Mittelalter. Vom griechischen Naturphilosophen DEMOKRIT (etwa 460–370 v. u. Z.), der als einer der Begründer der antiken Atomistik gilt, wird berichtet, daß er sich die folgende Frage vorgelegt habe: Durch fortlaufende Teilung eines Apfels erhält man zwei Hälften, vier Viertel usw. Gibt es eine Grenze der Teilbarkeit, wobei die kleinsten Teilchen nicht mehr die Eigenschaften des Apfels tragen? Er war der Meinung, daß es solche kleinsten, unteilbaren Teilchen gibt und nannte sie *Atome*. Diese seien unveränderlich, von einfacher geometrischer Gestalt und in unendlich vielen Sorten vorhanden. Alle komplizierteren Dinge wie Feuer, Wasser, Luft und Erde erklärte er als Wirbel der in der Leere bewegten Atome.

Zu Lebzeiten von DEMOKRIT waren in der *griechischen Naturphilosophie* schon verschiedene Lehren bekannt, die als Ursprung und Ende aller Dinge einzelne „Urelemente" annahmen. So bezeichnete THALES von Milet (etwa 625–546 v. u. Z.) das Wasser als Urelement, dagegen ANAXIMENES von Milet (etwa 585–525 v. u. Z.) die Luft und HERAKLIT von Ephesos (etwa 544–483 v. u. Z.) das Feuer. EMPEDOKLES aus Agrigent (etwa 495–435 v. u. Z.) begründete die *Vierelementelehre*, in der Feuer, Luft, Wasser und Erde die Rolle von Urelementen spielten. Die von DEMOKRIT aufgestellte Atomhypothese kann daher verstanden werden als konsequente Anwendung der Grundidee dieser Lehren: Die Vielfalt der komplizierten Dinge entsteht durch unterschiedliche Zusammenlagerung einfacherer Elemente.

In der *antiken Naturphilosophie* hat der Atombegriff nachweislich seinen Ursprung. Aber die moderne Atomistik entwickelte sich unabhängig von ihren antiken Vorgängern, da diese schnell in Vergessenheit gerieten. Viel wichtiger war das in der gesamten griechischen Naturphilosophie enthaltene Prinzip, wonach die Eigenschaften eines Ganzen auf die Eigenschaften seiner Teile zurückgeführt werden können. Dieses Prinzip ist bis in die moderne Physik hinein erfolgreich.

Im Mittelalter setzten sich unter dem Einfluß der katholischen Kirche in Europa die Lehren von ARISTOTELES aus Stagira (384–322 v. u. Z.) und PLATON (427–347 v. u. Z.) durch. Beide Philosophen waren Anhänger der Vierelementelehre. PLATON betrachtete die vier Elemente als Verwirklichungen von Ideen, während sie ARISTOTELES mit den vier Grundeigenschaften der Stoffe warm, kalt, trocken und feucht in Zusammenhang brachte. Den vier bekannten Elementen fügte ARISTOTELES mit dem Äther ein fünftes, eine *quinta essentia* hinzu. Die *Ätherhypothese* konnte erst durch die 1905 von EINSTEIN im Zusammenhang mit der speziellen Relativitätstheorie gegebene Interpretation des Michelsonversuchs widerlegt werden.

ARISTOTELES war ein vielseitiger Wissenschaftler von beeindruckender Geistesgröße. Er gab eine zusammenfassende Darstellung des gesamten Wissens seiner Zeit. In der Frage der Teilbarkeit der Dinge war er allerdings anderer Auffassung als die Atomisten. Die Schriften von DEMOKRIT wurden daher vernichtet, und seine Lehren sind uns nur durch Äußerungen anderer Autoren überliefert.

Obwohl die mittelalterliche Scholastik in ihrem Bemühen um Auslegungen und Deutungen der Lehren des ARISTOTELES anscheinend eine Stagnation der geistigen Entwicklung brachte, vollzog sich dennoch eine bedeutende Vervollkommnung der Experimentiertechniken. Es war die Zeit der *Alchemie*, die in Schwefel, Quecksilber und Salz Verwirklichungen der Eigenschaften Brennbarkeit, Flüchtigkeit und Löslichkeit erblickte und alle Stoffe auf diese drei Bestandteile zurückführen wollte. Sie wurde geprägt durch ihre mystische Komponente, die vor allem auf den byzantinischen, chinesischen und indischen Einfluß zurückzuführen war. Die Suche nach dem „Stein der Weisen", der das Leben verlängern und Metalle in Gold verwandeln sollte, trat in der Zeit ihrer stärksten Entfaltung, im arabischen Raum im 8.–11. Jahrhundert, in Mitteleuropa im 12.–14. Jahrhundert, in den Vordergrund. Besonders wichtig für die Entwicklung einer materialistischen Auffassung vom Leben war die These des PARACELSUS (Bombastus von Hohenheim, 1493–1541), daß auch lebende Organismen einschließlich des Menschen aus den drei Grundstoffen der Alchemie aufgebaut seien.

1.1.1.2. Erste quantitative chemische Gesetze. Die Anfänge der Naturwissenschaft sind gekennzeichnet durch konsequente Abkehr von der spekulativen Interpretation des Naturgeschehens und Hinwendung zum Experiment. Die arabischen Länder bildeten den Ausgangspunkt einer so orientierten wissenschaftlichen Tätigkeit. Zum Beispiel arbeiteten zur Zeit der Herrschaft des Kalifen HARUN AL RASCHID (786–809) in Bagdad etwa 800 Gelehrte verschiedener Nationalitäten und Glaubensbekenntnisse an vorwiegend praktisch orientierten Problemen. Anfänge einer Systematisierung der damals bekannten Stoffe gehen auf DSCHABIR (8. Jh.), vor allem aber auf AL-RAZI (um 865–925) und seine Schüler zurück. Da jegliche exakten Kenntnisse zur stofflichen Zusammensetzung noch fehlten, erfolgte die Einteilung allerdings nach äußerlichen Kriterien, wie Schmelzbarkeit, Glanz und Verdampfbarkeit der Stoffe. Die – oft schon richtig beobachteten – chemischen Reaktionen konnten naturgemäß auf solcher Grundlage nicht gedeutet werden.

In Europa forderte erstmals der englische Franziskanermönch BACON (etwa 1214–1294) eine neue Naturwissenschaft auf der Grundlage von Experiment und Mathematik. Seine Ideen ähnelten denen von LEONARDO DA VINCI (1452–1519).

In bezug auf praktische Probleme wie Baustatik und astronomische Zeitzyklen ging man schon in alten Kulturen von solchen Grundsätzen aus. Es galt, diese Prinzipien in alle Gebiete wissenschaftlicher Tätigkeit einzuführen. In diesem Sinne wirkten die Wiederentdecker der Atomidee, der französische Philosoph GASSENDI (1592–1655) und JUNGIUS (1587–1657), und andere bedeutende Wissenschaftler des 15. bis 17. Jahrhunderts, wie KOPERNIKUS (1473–1543), GALILEI (1564–1642) und KEPLER (1571–1630). Einer der ersten Kritiker der *Quecksilber-Schwefel-Salz-Lehre* der Alchimisten war der englische Chemiker BOYLE (1627–1691). In seinem 1661 erschienenen Buch „Der skeptische Chemiker" betrachtete er als *Elemente* nur die Stoffe, die chemisch einheitlich und nicht weiter zerlegbar waren. Kennzeichnend für die neue, wissenschaftliche Arbeitsweise ist die Entdeckung des *Boyle-Mariotteschen Gesetzes*, das 1662 von ihm, HOOKE (1635–1703) und TOWNLEY und 1676 unabhängig davon vom französischen Physiker MARIOTTE (1620–1684) gefunden wurde. Danach ist das Produkt aus zusammengehörigen Werten von Druck und Volumen eines Gases bei gleichbleibender Temperatur konstant. BOYLE schloß daraus, daß sich die Atome im leeren Raum mit großen Geschwindigkeiten bewegen und der Druck auf die Gefäßwände Ergebnis ihrer Stöße ist.

Erste deutliche Hinweise auf die Existenz von Atomen mußten aus der Chemie hervorgehen. Denn gerade hier werden Vorgänge untersucht, bei denen die gemessenen makroskopischen Mengen von Anfangs- und Endprodukten einer Reaktion Abbild des Verhältnisses sind, in dem sich im Mikroskopischen unvorstellbar häufig und in immer der gleichen Weise Atome zu Molekülen verbinden.

Noch fehlte allerdings das qualitative Verständnis von Reaktionsabläufen als Vorbedingung für deren quantitative Behandlung. Hier erwies sich die vor allem auf STAHL (1660–1734) zurückgehende Phlogistonlehre als heuristisch wertvoll. Danach sollte aus Metallen und anderen Stoffen beim Verbrennen ein „Feuerstoff", das Phlogiston, entweichen. Wenngleich diese Hypothese die im Ansatz schon vorhandenen richtigen Vorstellungen über Verbindungen und Elemente vorübergehend in Frage stellte, führte sie durch die Beschreibung zahlreicher Oxydations- und Reduktionsvorgänge zum Verständnis für den Zusammenhang zwischen Hin- und Rückreaktion bei chemischen Umsetzungen.

Nach dem Tode von LAVOISIER gewinnen folgerichtig Versuche zur Aufdeckung quantitativer chemischer Gesetzmäßigkeiten an Bedeutung. So fand der deutsche Chemiker RICHTER (1762–1807), daß sich Säuren und Basen immer nur in ganz bestimmten Massenverhältnissen neutralisieren, wobei Salze entstehen. Der französische Chemiker PROUST (1754–1826) formulierte 1801 als Ergebnis seiner Experimente zu Sauerstoff-, Schwefel- und Chlorverbindungen von Metallen das *Gesetz der kon-*

stanten Proportionen. Nach diesem Gesetz enthalten chemische Verbindungen ihre Bestandteile immer in konstanten Verhältnissen. Der englische Physiker DALTON (1766–1844) war wie NEWTON (1643–1727) ein überzeugter Anhänger der Atomidee von GASSENDI (1592 bis 1655). Er suchte nach Beweisen für die Existenz der Atome und wurde so zum Begründer der *wissenschaftlichen Atomistik.* Im Jahre 1805 formulierte er das *Gesetz der multiplen Proportionen.* Darin kommt die Erfahrung zum Ausdruck, daß sich die Mengen eines Elements, die sich mit einer gegebenen Menge eines anderen Elements verbinden, wie kleine ganze Zahlen verhalten. Zum Beispiel verbinden sich in den verschiedenen Stickoxiden 10 g Stickstoff mit 5,7 g/11,4 g/17,1 g/22,8 g oder 28,5 g Sauerstoff. Die unterschiedlichen Massen des Sauerstoffs verhalten sich also wie 1 : 2 : 3 : 4 : 5, und so müssen sich auch die Anzahlen der Sauerstoffatome verhalten, die pro Stickstoffatom in den Stickoxidmolekülen gebunden sind. Heute drücken wir diese Tatsache durch die chemischen Symbole der Moleküle aus: N_2O, NO, N_2O_3, NO_2, N_2O_5.

Für DALTON war ein Element ein Stoff, der aus einer Sorte von „*Elementaratomen*" bestand. Davon unterschied er die „*zusammengesetzten Atome*" der Verbindungen, die heutigen Moleküle. Er hielt es für sehr wichtig, eine Tabelle der *relativen Atommassen* dieser verschiedenen Atome aufzustellen. In den 1803 und 1805 veröffentlichten Tabellen (Tab. 1.1.1-1) ordnete er dem Wasserstoff die relative Atommasse 1 zu.

Tab. 1.1.1-1 Chemische Symbole und Atom- bzw. Molekularmassen von DALTON, 1808/10; zum Vergleich sind die heutigen Werte für die entsprechenden Formeleinheiten angegeben (gerundet)

	Symbol	Dalton	heute	
Wasserstoff	⊙	1	1	
Stickstoff	⦶	5	14	
Kohlenstoff	●	5	12	
Sauerstoff	○	7	16	
Phosphor		9	31	
Schwefel	⊕	13	32	
„Kalk" (Ca)	◎	23	40	
„Natron" (Na)		28	23	
Quecksilber		84 oder 168	200	
Kupfer	©	56 oder 28	64	
Wasser (H + O)	⊙○	8	18	H_2O
Ammoniak (N + H)	⊙⦶	6	17	NH_3
„ölbildendes Gas" (C + H) (Ethylen)	⊙●	6	28	C_2H_4
Kohlensäure (C + 20)	○●○	19	44	CO_2
Schwefelsäure (S + 30)		34	80	SO_3
Schwefelwasserstoff (S + 3 H)		16	34	H_2S
Essigsäure (2 C + 2 „Wasser")		26	60	CH_3COOH
Alkohol (3 C + H)		16	46	C_2H_5OH
Zucker (CO_2 + „Alkohol")		35	342	$C_{12}H_{22}O_{11}$

Die Tabellen enthielten aber noch erhebliche Fehler, denn DALTON ging von z. T. unrichtigen Vorstellungen über die Zusammensetzung einiger Moleküle aus; so nahm er auch an, daß Elemente wie Wasserstoff, Stickstoff und Sauerstoff aus einzelnen Atomen aufgebaut seien.
Heute wird für Physik und Chemie einheitlich als Basis der relativen Atommasse 1/12 der Masse eines Atoms des Kohlenstoffisotops $^{12}_{6}C$ benutzt.

1.1.1.3. Avogadrosche Hypothese. Daltons relative Atommassen wurden durch Wägung der Anfangs- und Endprodukte chemischer Reaktionen bestimmt. Dabei bestanden Unklarheiten über die Anzahl der darin enthaltenen Teilchen. Die 1811 vom italienischen Physiker AVOGADRO (1776–1856) aufgestellte Hypothese trug wesentlich zur Behebung dieser Unklarheiten bei. AVOGADRO nahm an, daß gleiche Volumina verschiedener Gase unter gleichen Bedingungen des Drucks und der Temperatur gleich viel Teilchen enthalten. Diese wichtige Aussage entstand in einer Epoche umfassenden Studiums der Physik und Chemie von Gasen und Dämpfen. Hervorragende Ergebnisse sind auch das 1803 von DALTON entdeckte *Gesetz der Partialdrücke* und das *Gay-Lussacsche Gesetz*. Das Erstgenannte besagt, daß der Gesamtdruck eines Gasgemischs gleich der Summe der Einzeldrücke seiner Bestandteile ist. Das Gay-Lussacsche Gesetz beschreibt die Volumenausdehnung von Gasen und Dämpfen unter konstantem Druck. Nach seiner ursprünglichen Formulierung nimmt das Gasvolumen bei einer Temperaturerhöhung um 100 K um das 0,375-fache zu. Heute gilt für den thermischen Ausdehnungskoeffizienten eines verdünnten Gases der Wert $0{,}00366\ K^{-1} \approx 1/273{,}16\ K$.
Ausgangspunkt der Hypothese des AVOGADRO waren Experimente von GAY-LUSSAC (1778 bis 1850). Dieser untersuchte bei chemischen Reaktionen zwischen Gasen die Volumina der Anfangs- und Endprodukte für feste Werte des Drucks und der Temperatur. Zum Beispiel fand er, daß sich 1 Raumteil (RT) Chlor und 1 RT Wasserstoff zu 2 RT Chlorwasserstoff verbinden. Nach AVOGADRO entstehen aus N Teilchen Chlor und N Teilchen Wasserstoff $2N$ Teilchen Chlorwasserstoff. Folglich müssen sich die Teilchen von Chlor und Wasserstoff während der Reaktion geteilt haben. Heute wissen wir, daß Chlor und Wasserstoff unter normalen Bedingungen aus Molekülen zu je zwei Atomen bestehen.
Nach der Auffassung von AVOGADRO gab es drei Sorten kleinster Teilchen: 1. die „*Elementarmoleküle*“ (molécules elémentaires) der nicht weiter zerlegbaren Elemente, wozu er alle Metalle zählte, – 2. die „*integrierenden Moleküle*“ (molécules intégrantes) der zusammengesetzten Elemente, hierzu rechnete er die meisten Nichtmetalle und Gase, – 3. die „*konstituierenden Moleküle*“ (molécules constituentes) der Verbindungen. Atome im heutigen Sinne sind nur die unter 1. genannten Teilchen.
Die Avogadrosche Hypothese ermöglichte eine genaue Bestimmung der atomaren Bestandteile von Molekülen, was er selbst am Beispiel von Wasser und Ammoniak zeigte. Sie hatte damit eine deutliche Verbesserung der Tabellen von relativen Atom- und Molekülmassen zur Folge. Obwohl BERZELIUS (1779–1848) schon 1818 eine solche Tabelle von hoher Genauigkeit für etwa 2000 Verbindungen vorlegte, geriet die Avogadrosche Hypothese wieder in Vergessenheit. Es ist das Verdienst des italienischen Chemikers CANNIZZARO (1826–1910), ihre Nützlichkeit 1856 endgültig bewiesen zu haben.
Genauere Bestimmungen der auf Wasserstoff bezogenen relativen Atommassen lieferten häufig nahezu ganzzahlige Werte. Das veranlaßte den englischen Arzt PROUT (1785–1850), das Wasserstoffatom als elementaren Bestandteil aller Atome anzunehmen. Erst im Verlauf der Erforschung der Struktur des Atoms im 20. Jahrhundert konnte der Wahrheitsgehalt dieser damals spekulativ geäußerten Ansicht freigelegt werden.

1.1.1.4. Periodensystem der Elemente. Das schnelle Wachstum der Industrieproduktion ausgangs des 18. und vor allem im 19. Jahrhundert brachte einen raschen Aufschwung der chemisch-physikalischen Forschung. Die vordringliche Aufgabe bestand daher in der Systematisierung der Einzelergebnisse mit dem Ziel, die Gesetzmäßigkeiten der Verwandtschaft des chemischen und physikalischen Verhaltens der Elemente aufzudecken.
DOEBEREINER fand 1817, daß die relativen Atommassen von je drei chemisch verwandten Elementen ziemlich genau arithmetische Folgen bilden. Er nannte diese Elementefolgen *Triaden*, z. B. Chlor 35,5 – Brom 80 – Iod 127 oder Calcium 40 – Strontium 88 – Barium 137. PETTENKOFER zeigte, daß chemische Verwandtschaften auch in Gruppen von mehr als drei Elementen bestehen, z. B. Fluor – Chlor – Brom – Iod und Magnesium – Calcium – Strontium – Barium. Der englische Chemiker NEWLANDS stellte 1864 fest, daß in einer nach steigenden relativen Atommassen geordneten Reihe der etwa 60 damals bekannten Elemente nach jeweils sieben Elementen ähnliche chemische Eigenschaften wiederkehrten. Er nannte diese Gruppen *Oktaven*.
Den endgültigen Durchbruch zu einer umfassenden Systematik der Elemente erzielten zwei Wissenschaftler unabhängig voneinander etwa zur gleichen Zeit. L. MEYER und MENDELEJEW gelangten ausgehend von einer Anordnung der 63 damals bekannten Elemente nach steigenden

relativen Atommassen in den Jahren 1868–1871 zu nahezu identischen Formulierungen für das *Periodensystem der Elemente* (Tafel 80).
L. Meyer untersuchte die *Änderung der chemischen und physikalischen Eigenschaften* in der Elementenreihe. Dabei fand er eine deutliche *Periodizität der Molvolumina der Atome*. Als *Molvolumen* wird der Quotient aus Molmasse und Dichte eines Elements im festen oder flüssigen Zustand bezeichnet. Es ist somit proportional dem einem Atom durchschnittlich zur Verfügung stehenden Volumen. Die Alkalimetalle (Li, Na, K, Rb, Cs) bilden die Maxima, während die Erdalkalimetalle (Be, Mg, Ca, Sr, Ba) in den stark abfallenden und die Halogene (F, Cl, Br, I) in den aufsteigenden Kurventeilen liegen.
Mendelejew ging von den in der Elementenreihe *regelmäßig wiederkehrenden äquivalenten chemischen Eigenschaften* aus und ordnete solche Elemente untereinander an, die gleichartige Verbindungen mit Wasserstoff, Sauerstoff und den Halogenen bilden. Die so entstandene Tabelle enthält, die später entdeckten Edelgase eingeschlossen, 8 Gruppen jeweils chemisch ähnlicher Elemente.
Beide Darstellungen wiesen Lücken auf für noch zu entdeckende Elemente, deren physikalische und chemische Eigenschaften aus ihren Stellungen im Periodensystem abgeleitet werden konnten. Am Beispiel des von Mendelejew als Eka-Silicium bezeichneten Germanium, das 1886 von C. Winkler entdeckt wurde, vermittelt Tab. 1.1.1-2 einen Eindruck hiervon.

Tab. 1.1.1-2 Einige Eigenschaften von Germanium (Eka-Silicium)

	vorausgesagt Mendelejew (1871)	bestimmt Winkler (1885/86)
Relative Atommasse	72	72,6
Dichte in gcm^{-3}	5,5	5,35
Oxid	EsO_2	GeO_2
Dichte in gcm^{-3}	4,7	4,70
Chlorid	$EsCl_4$ flüssig	$GeCl_4$ flüssig
Dichte in gcm^{-3}	1,9	1,887
Siedetemperatur in K	etwa 363	356

Auch die 1875 von de Boisbaudran bzw. 1879 von Nilson gefundenen Elemente Gallium und Scandium stimmten in ihren Eigenschaften mit den Voraussagen von Mendelejew überein.
Moderne Darstellungen des Periodensystems unterscheiden sich vor allem in einer Hinsicht von ihren Vorgängern nach L. Meyer und Mendelejew: Der in der natürlichen Elementenreihe monoton wachsende Parameter ist nicht die relative Atommasse, sondern die Kernladungszahl. Eine entsprechende Vermutung wurde aber erst 1911 durch van den Broek geäußert und 1913 durch Moseley belegt, der die Abhängigkeit der Frequenz einer Spektrallinie der charakteristischen Röntgenstrahlung der Elemente von deren Ordnungs- (= Kernladungs-) zahl feststellte (↑ 1.1.4.7.). Im Rahmen der Mendelejew und L. Meyer bekannten Elemente wäre das auch nur bezüglich der Reihenfolge von Tellur und Iod sowie Cobalt und Nickel bedeutsam gewesen. Diese wurde aufgrund der chemischen Eigenschaften richtig festgelegt.
Zu den Ursachen der im Periodensystem zum Ausdruck kommenden Gesetzmäßigkeiten konnten damals nur Vermutungen geäußert werden. Einige Wissenschaftler waren der richtigen Meinung, daß sie im unterschiedlichen Aufbau der Atome aus Teilchen einer neuen Ebene der Struktur der Materie liegen.

1.1.1.5. Erste absolute Angaben über das Atom. Einen ersten anschaulichen Hinweis auf die Existenz von Atomen fand 1827 der schottische Botaniker Brown. Unter dem Mikroskop beobachtete er, daß im Wasser fein verteilter Blütenstaub unregelmäßige Bewegungen ausführte. Er nahm allerdings irrtümlich an, daß organische Moleküle eine entsprechende aktive Bewegungsfähigkeit besitzen.
Die Ursachen der *Brownschen Bewegung*, wie dieser Effekt heute genannt wird, sind die Zusammenstöße der Flüssigkeitsmoleküle mit den Pollen. Wegen der geringen Größe der Pollen entstehen dabei resultierende Kraftstöße, die darüber hinaus intensiv genug sind, um die Pollen in zitternde Diffusionsbewegungen zu versetzen. Diese Erklärung fand 1888 der Franzose Gouy, nachdem der Effekt lange Zeit für ein Analogon zur Bewegung von Staub im Sonnenlicht und damit für uninteressant gehalten wurde.
Die quantitative Beschreibung der Brownschen Bewegung wurde 1905 durch Einstein und 1906 unabhängig von ihm durch den polnischen Physiker von Smoluchowski gegeben. In beiden Arbeiten wurde die Bewegung der Teilchen einer Suspension auf die Wärmebewegung der Flüssigkeitsmoleküle zurückgeführt. Perrin lieferte 1909 einen Beweis für die Richtigkeit dieser Theorie, indem er die im Zusammenwirken von Schwerkraft und Brownscher Diffusionsbewegung entstehende Dichteverteilung der Teilchen einer Suspension untersuchte. Diese Arbeiten trugen entscheidend zur endgültigen Anerkennung der Atomhypothese bei.
Das Bestreben, die Atome durch Ausdehnung und Masse zu charakterisieren, bestand schon unter den frühen Anhängern der Atomistik

So schätzte FRANKLIN 1773 die Ausdehnung der Atome aus der Dicke einer Ölschicht auf Wasser ab und erhielt einige 10^{-9} m. Erste verläßliche Angaben zu den Atommassen und den Ausdehungen der Atome machte im Jahre 1865 der österreichische Physiker LOSCHMIDT. Er untersuchte die wechselseitige Diffusion von Gasen, wobei er Ergebnisse der kinetischen Gastheorie benutzte. LOSCHMIDT zeigte, daß die Durchmesser der Atome die Größenordnung 10^{-10} m besitzen und daß benachbarte Atome eines Gases unter Normalbedingungen einen etwa 10mal so großen Abstand haben. Für die Masse eines Wasserstoffatoms gab er etwa 10^{-27} kg an. Er errechnete, daß 1 cm^3 eines beliebigen Gases bei normalem Atmosphärendruck und 0 °C etwa $2 \cdot 10^{19}$ Moleküle enthält. Damit bestätigte er die Avogadrosche Hypothese. Loschmidts Wert war etwas zu klein. Heute bezeichnet man als *Avogadrokonstante* $N_A = (6{,}022045 \pm 0{,}000031) \cdot 10^{23}$ mol^{-1} die Anzahl der in einem Mol, entsprechend 0,012 kg des Kohlenstoffisotops $^{12}_{6}C$, enthaltenen Atome.

1912 kannte man bereits mehr als zehn voneinander unabhängige Verfahren zur Bestimmung der Avogadrokonstante, die im deutschen Sprachgebrauch auch als Loschmidtsche Zahl bezeichnet wird.

1.1.1.6. Atomistik in der Physik des 19. Jahrhunderts. Ein Ziel der Atomistik ist es, Materialeigenschaften sowie makroskopische Vorgänge auf atomare Eigenschaften bzw. Prozesse zwischen Atomen zurückzuführen. Zu den ersten bedeutenden Erfolgen auf physikalischem Gebiet zählt die Erforschung der *Natur der Wärme*. Noch im 18. Jahrhundert wurden Wärme und Kälte durch bestimmte zu- und abführbare Stoffe erklärt. Dagegen war LOMONOSSOW der Meinung, daß Wärme in einer kreisenden Bewegung der kleinsten Teilchen der Stoffe besteht. Die innere Energie oder Wärmeenergie eines Körpers ist die Summe der mechanischen Energien seiner Atome.

Das *Prinzip der Äquivalenz von Wärme und mechanischer Energie* wurde erstmals von J. R. MAYER im Jahre 1842 präzise formuliert. Den Ausgangspunkt seiner Überlegungen bildete die Tatsache, daß für eine bestimmte Erwärmung eines Gases bei konstantem Druck mehr Energie als bei konstantem Volumen aufgewendet werden muß. JOULE bestimmte genaue Werte für das mechanische und das elektrische Wärmeäquivalent, und HELMHOLTZ verallgemeinerte 1847 diese grundlegende Erkenntnis zum *Satz von der Erhaltung der Energie:*

In einem abgeschlossenen System ist die Summe aller Energien konstant, es treten nur Umwandlungen von einer Energieform in andere auf.

Für die Atomistik ist dieser Satz von großer Bedeutung, denn er bringt zum Ausdruck, daß atomare und makroskopische Phänomene einheitlichen physikalischen Gesetzen unterliegen.

Die quantitative Erklärung von physikalischen Eigenschaften eines Systems durch eine atomistische Theorie gelang erstmals in der *kinetischen Gastheorie*. Diese nimmt an, daß sich die Teilchen des betrachteten Systems, abgesehen von kurzzeitigen elastischen Stößen, kräftefrei bewegen. Sie ist daher besonders für verdünnte Gase geeignet. Begründer dieser Theorie ist BERNOULLI. In seinem 1738 erschienenen Werk „Hydrodynamica" leitet er aus gaskinetischen Überlegungen eine heute nach ihm benannte Grundgleichung der Strömungslehre her. Trotz einzelner weiterführender Beiträge fand diese Betrachtungsweise lange keine Anerkennung. Die Grundgedanken wurden 1856 bzw. 1857 von KÖNIG und CLAUSIUS wiederentdeckt, und MAXWELL und BOLTZMANN schlossen in ihren Arbeiten aus den Jahren 1860 bzw. 1894 den Ausbau der Grundlagen der kinetischen Gastheorie im wesentlichen ab. Die Theorie geht davon aus, daß jeder beliebig komplizierte Zustand des Systems durch Angabe der Orte und Geschwindigkeiten seiner Teilchen mikroskopisch beschrieben werden kann. Dieser Mikrozustand unterliegt zeitlichen Änderungen, denn die Teilchen stoßen häufig miteinander zusammen, z. B. stößt in Wasserstoff unter Normalbedingungen jedes Atom in einer Sekunde auf $1{,}5 \cdot 10^{10}$ andere Atome. Im Ergebnis dieser Teilchenwechselwirkung entsteht aus jedem beliebigen Anfangszustand nach hinreichend langer Zeit ein Zustand, dessen zeitlich mittleres Verhalten durch wenige makroskopische Parameter wie Volumen V, Druck p, Temperatur T usw. charakterisiert ist. Dieser Zustand wird als *thermisches Gleichgewicht* bezeichnet. Falls keine äußeren Kräfte, wie die Schwerkraft, wirken, ist das thermische Gleichgewicht mikroskopisch dadurch gekennzeichnet, daß die Teilchen im Zeitmittel homogen im verfügbaren Volumen verteilt sind und eine isotrope Maxwellsche Geschwindigkeitsverteilung besitzen.

Besonders deutlich wird die Leistungsfähigkeit der kinetischen Gastheorie bei der Beschreibung des thermodynamischen Verhaltens verdünnter Gase. Der *Gasdruck* auf die Gefäßwand entsteht dynamisch, indem die auftreffenden Teilchen die Normalkomponenten ihrer Impulse bei Reflexion umkehren und dabei Kraftstöße auf die Wand ausüben. Summation über alle Teilchen liefert $p = (2/3)\, n\bar{\varepsilon}$, wobei n die Dichte der Teilchen (Teilchenzahl pro Volumen) und $\bar{\varepsilon}$ ihre mittlere kinetische Energie bedeuten. Entsprechende atomistische Erklärungen konnten für Diffusion und Wärmeleitung sowie für Schallausbreitung und andere Eigenschaften von Gasen gegeben werden.

1.1.2. Freie Elektronen und Ionen

Ausgehend von ihren chemischen Erfahrungen vermuteten schon die frühen Atomisten des 19. Jahrhunderts, daß die Atome unterschiedlicher Elemente aus unterschiedlich vielen einfacheren Teilchen aufgebaut sind. Wir wissen heute, daß Elektron, Proton und Neutron Bausteine der Atome sind. Aufklärung über die Bestandteile des Atoms konnte nur die Untersuchung zerstörter Atome bringen. Solche „zerstörten" Atome werden als *Ionen* bezeichnet. Sie wurden erstmals bei der *Elektrolyse* wäßriger Lösungen nachgewiesen. Dabei versteht man unter Elektrolyse die bei Stromfluß durch die Lösung an den eingetauchten Elektroden ablaufende Zersetzung der gelösten Substanz. Zwar brachte das Studium der Elektrolyse die Entdeckung der Elementarladung, die Elektronen selbst als Träger der Elementarladung traten aber nicht in Erscheinung.

Dagegen eröffnete die Untersuchung des Stromdurchgangs durch verdünnte Gase mit der Entdeckung des Elektrons die Forschung zur Struktur des Atoms einerseits und andererseits die Entwicklung ganzer Industriezweige zur Herstellung elektronenphysikalischer Geräte.

1.1.2.1. Entdeckung der Elementarladung. Aufgrund seiner Experimente formulierte FARADAY 1834 die beiden *Grundgesetze der Elektrolyse*. Diese besagen, daß die abgeschiedenen Stoffmengen den transportierten Ladungsmengen proportional sind und daß durch die Ladungsmenge 96493 C gerade $1/n$ Mol eines n-wertigen Stoffs abgeschieden werden. Die Anzahl der Atome eines Mols ist durch die Avogadrokonstante (↑ 1.1.1.5.) gegeben. Folglich transportiert jedes einwertige Ion die Ladungsmenge $e = N_A^{-1} \cdot 96493\ \text{C} = (1{,}6020 \pm 0{,}0004) \cdot 10^{-19}\ \text{C}$. Eine kleinere Ionenladung fand man nicht; mehrwertige Ionen tragen ein ganzzahliges Vielfaches dieser *Elementarladung*. So erscheint es konsequent, wenn HELMHOLTZ diese Ladungsmenge 1881 als das Atom der Elektrizität bezeichnete. STONEY prägte dafür den Begriff *Elektron*.

Die Zerlegung von Atomen in wäßriger Lösung konnte erst 1887 durch die Dissoziationstheorie des schwedischen Chemikers ARRHENIUS zufriedenstellend erklärt werden. FARADAY war der Meinung, daß diese Zerlegung in Ionen erst mit dem Anlegen einer elektrischen Spannung einsetzt.

1.1.2.2. Katoden- und Kanalstrahlen. Schon 1838 wurden von FARADAY Leuchterscheinungen beim Stromdurchgang durch verdünnte Gase beobachtet. MASSON untersuchte 1853 den Ladungstransport durch einen stark evakuierten Raum und fand ein phosphoreszierendes Leuchten. Die eigentliche Untersuchung dieser Effekte begann jedoch erst 1858 mit den Arbeiten von PLÜCKER sowie seiner Schüler GOLDSTEIN und HITTORF.

PLÜCKER benutzte Gasentladungsröhren mit heißer Katode und einem Gasdruck von etwa 10^2 Pa. Bei Stromdurchgang beginnt vor allem der Raum vor der Anode in einer für das Gas charakteristischen Farbe zu leuchten. Die Leuchterscheinung verändert bei Druckerniedrigung ihre Struktur und verringert sich insgesamt, während die Stromstärke zunimmt. Schließlich verschwindet das Gasleuchten bei hohem Vakuum ganz, und an seine Stelle tritt eine Lumineszenz der gegenüber der Katode liegenden Glaswand. Dieses Leuchten wird von einer Strahlung hervorgerufen, die GOLDSTEIN 1876 als *Katodenstrahlung* bezeichnete.

Von den zahlreichen Beiträgen zur Erforschung der Katodenstrahlen sind die des englischen Physikers CROOKES hervorzuheben. 1878 gab er vor der Londoner Königlichen Akademie einen Bericht zum Stand der Katodenstrahlforschung. CROOKES erklärte, daß Katodenstrahlen ein Strom schnell bewegter, negativ geladener Teilchen seien, deren Ausdehnung deutlich unter der eines Atoms liege. Er erkannte die Bedeutung dieser neuen Existenzform der Materie und bezeichnete sie als vierten Aggregatzustand sowie, in Anlehnung an FARADAY, als „strahlende Materie". Nachdem H. HERTZ 1892 gezeigt hatte, daß Katodenstrahlen Metallfolien durchdringen, kam es zu einer ausgedehnten Untersuchung ihrer Wechselwirkung mit Körpern außerhalb des Vakuums durch LENARD.

Die abschließende Klärung der Natur der Katodenstrahlung gelang THOMSON im Jahre 1897, indem er die spezifische Ladung der Teilchen der Katodenstrahlung ermittelte. Als *spezifische Ladung* wird der Quotient aus Ladung und Masse der Teilchen bezeichnet. Entsprechend der 1881 von STONEY ausgesprochenen Hypothese, daß es sich bei den Teilchen der Katodenstrahlung um die Elektrizitätsatome, also um die Elektronen, handele, erhielt man aus spezifischer und Elementarladung für die Elektronenmasse etwa 1/2000 der Masse des Wasserstoffatoms.

Es war naheliegend, die Atome des verdünnten Gases für den Ursprung der Elektronen zu halten. Dann mußten aber auch die zerstörten Atome, positiv geladene Ionen, nachweisbar sein. Sie wurden 1886 von GOLDBERG entdeckt und als *Kanalstrahlen* bezeichnet. Kanalstrahlen treten hinter der Katode auf einem Leuchtschirm in Erscheinung, wenn man den in Richtung Katode strömenden positiven Ionen durch eine Öffnung in der Katode den Austritt aus dem Entladungsraum ermöglicht. Nach diesem Prinzip arbeiteten die ersten Ionenstrahlquellen, die eine notwendige Voraussetzung für die

Entwicklung der Massenspektroskopie darstellten.

1.1.2.3. Der Weg zur Massenspektroskopie. 1910 fanden SODDY und MARCKWALD bei der Untersuchung radioaktiver Zerfallsreihen († 1.1.5.4.), daß sich die beiden hinsichtlich ihrer radioaktiven Eigenschaften deutlich verschiedenen Substanzen Radium und Mesothor 1 chemisch nicht unterscheiden. Sie nannten diese Atomsorten *Isotope* von Radium. Die damals üblichen chemischen Verfahren zur Bestimmung der relativen Atommassen konnten offensichtlich nur Mittelwerte über das jeweils untersuchte Isotopengemisch liefern. Daher mußten physikalische Verfahren gefunden werden, die eine direkte Bestimmung der Teilchenmasse ermöglichen. Solche Verfahren wurden im Zusammenhang mit der Untersuchung von Katoden- und Kanalstrahlen entwickelt. Ihre physikalische Grundlage bildet die Kraftwirkung elektrischer und magnetischer Felder auf geladene Teilchen.

Die Bestimmung der spezifischen Ladung e/m des Elektrons durch THOMSON stellt die erste konsequente Nutzung dieses Prinzips dar. Dabei werden die Ablenkungen x und y der Elektronen in einem homogenen elektrischen bzw. magnetischen Feld senkrecht zur Strahlrichtung ermittelt. Die Bestimmung von e/m und der Elektronengeschwindigkeit v gelingt, da die Ablenkungen auf unterschiedliche Weise von den gesuchten Größen abhängen: $x = A_1 e/mv^2$, $y = A_2 e/mv$. Dabei sind A_1 und A_2 Gerätekonstanten.

Dieses Prinzip, das unter dem Namen *Parabelmethode* bekannt geworden ist, benutzte THOMSON 1911 zur Konstruktion des ersten *Massenspektrographen*. Ein aus einfach positiv geladenen Ionen unterschiedlicher Geschwindigkeiten bestehendes Kanalstrahlbündel wird gleichzeitig durch beide Felder beeinflußt. Sind die Felder gleichgerichtet, so bilden die Auftreffpunkte von einheitlich geladenen Ionen gleicher Masse, aber unterschiedlicher Geschwindigkeiten auf dem Schirm Parabeläste, wie aus den oben angegebenen Formeln folgt. Am Beispiel von Neon gelang THOMSON 1913 der Nachweis, daß auch nichtradioaktive Elemente Isotope besitzen.

Die Leistungsfähigkeit eines Massenspektrographen wird vorrangig durch sein Auflösungsvermögen bestimmt, d. h. durch die kleinste noch trennbare Differenz der relativen Ionenmassen. Moderne Geräte erreichen Auflösungen von 10^{-6} der Masse des Wasserstoffatoms, was durch schrittweise Überwindung der Nachteile der Parabelmethode erreicht wurde. Markante Punkte dieser Entwicklung stellen die 1918 von DEMPSTER, 1919 von ASTON und 1934 von MATTAUCH und HERZOG konstruierten Geräte dar.

1.1.2.4. Millikanversuch. Im Jahre 1909 führte der amerikanische Physiker MILLIKAN nach der von ihm entwickelten Öltröpfchenmethode erstmals eine direkte *Bestimmung der Elementarladung* durch. Er wiederholte das Experiment bis 1940 mehrmals mit wachsender Genauigkeit. MILLIKAN zeigte, daß die Ladung der beim Zerstäuben von Öl entstehenden Öltröpfchen vom Durchmesser 10^{-6} bis 10^{-7} m immer ein ganzzahliges Vielfaches der Elementarladung ist.

Das Meßprinzip des von ihm benutzten Schwebekondensators beruht darauf, daß Richtung und Betrag der Geschwindigkeit eines durch ein Fernrohr beobachteten Öltröpfchens durch das Gleichgewicht dreier Kräfte bestimmt sind: der Schwerkraft, der elektrischen Kraft sowie der bremsenden Kraft der Luftreibung. Die Ladung des Tröpfchens kann so aus seiner Fall- und Steiggeschwindigkeit ohne bzw. mit eingeschaltetem elektrischem Feld berechnet werden. Die unsichere Kenntnis der Reibung Tropfen – Luft bildet die Hauptfehlerquelle des Verfahrens. Aus genaueren Messungen der Elementarladung mit dieser Methode ergibt sich der Mittelwert $e = (1{,}604 \pm 0{,}005) \cdot 10^{-19}$ C. Er ist weniger genau als der aus Faraday- und Avogadrokonstante bestimmte Wert.

1.1.2.5. Elektronenspin und Stern-Gerlach-Versuch. Im Ergebnis der durchgeführten Experimente konnte das Elektron durch Masse und Ladung charakterisiert werden. Es erscheint verständlich, daß man ihm auch die Eigenschaft „Ausdehnung" zuerkannte und es somit zunächst als eine sehr kleine geladene Kugel betrachtete. Eine Möglichkeit zur Definition eines „Elektronenradius" liefert das Einsteinsche Konzept der Trägheit der Energie, $E = mc^2$ († 2.1.3.6.). Danach kann aus der bekannten Ruhmasse m des Elektrons auf die zur „Herstellung" eines Elektrons notwendige Energie geschlossen werden. Berücksichtigt man dabei nur die bei Konzentration der Elementarladung e innerhalb einer Kugel vom Radius r_0 gegen die elektrostatische Abstoßung zu verrichtende Arbeit, so erhält man den *klassischen Elektronenradius* $r_0 = e^2/4\pi\varepsilon_0 mc^2 \approx 2{,}818 \cdot 10^{-15}$ m. Nach der Mechanik konnte das ausgedehnte Elektron dann unabhängig von seiner Translationsbewegung einen Eigendrehimpuls, einen *Spin* († 2.1.3.9.) besitzen. Darüber hinaus stellte es als rotierende Ladungsverteilung einen Kreisstrom dar, dem nach der Elektrodynamik ein mit dem Spin $\boldsymbol{s}$ verknüpftes magnetisches Dipolmoment $\boldsymbol{\mu}$ († 2.2.2.5.) zukam. Ausgehend von solchen Überlegungen formulierten GOUDSMID und UHLENBECK 1925 ihre *Spinhypothese*, wonach dem Elektron neben Masse und Ladung auch ein Spin und ein magnetisches Dipolmoment zukommen müsse.

Quantenmechanisch ist die eingangs erläuterte Modellvorstellung nicht haltbar, da sie der

Physik makroskopischer Körper entnommen ist. Tatsächlich widerspricht das „Kugel-Elektron" auch in anderer Hinsicht elementaren Erfahrungen, denn die Rotationsgeschwindigkeit an seinem Äquator würde etwa 100mal größer als die Lichtgeschwindigkeit sein. Die Quantentheorie zeigt, daß der Spin auch unabhängig vom klassischen Modell eingeführt werden kann.
Einen ersten Hinweis auf die Existenz des Elektronenspins lieferte der *Stern-Gerlach-Versuch.* Die beiden deutschen Physiker untersuchten die Ablenkung von Atomstrahlen in inhomogenen Magnetfeldern erstmals 1922. Das Meßprinzip beruht darauf, daß auf ein

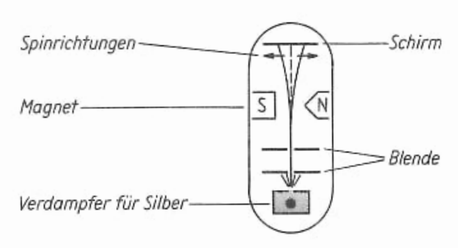


Abb. 1.1.2-1 Stern-Gerlach-Versuch

Atom je nach der Orientierung seines magnetischen Dipolmoments $\boldsymbol{\mu}$ in bezug auf das Magnetfeld $\boldsymbol{B}$ eine Kraft in Richtung wachsender ($\boldsymbol{\mu B} > 0$) oder abnehmender ($\boldsymbol{\mu B} < 0$) Feldstärke wirkt. Das historische Experiment wurde mit einem Strahl von Silberatomen im Zustand kleinster innerer Energie durchgeführt (Abb. 1.1.2-1). In diesem Zustand liefert die Bahnbewegung der Elektronen keinen Beitrag zum magnetischen Dipolmoment des Atoms. Die im Experiment beobachtete Aufspaltung des Atomstrahls in zwei Bündel symmetrisch zur Primärstrahlrichtung ist daher ein Beweis für die Existenz des Spinmagnetismus. Ein Strahl von einfach positiv geladenen Silberionen zeigt keine Aufspaltung, folglich ist das magnetische Dipolmoment des Silberatoms im Zustand kleinster innerer Energie gleich dem eines Elektrons. Diese Schlußfolgerung wird gestützt durch entsprechende Ergebnisse, die PHILLIPS und TAYLOR 1927 bei Durchführung des Experiments mit Wasserstoffatomen erhielten.
Die beobachtete Aufspaltung läßt nur *eine* Interpretation zu: Für die Feldkomponente μ_z des Vektors μ gibt es genau zwei mögliche Werte, die sich nur durch das Vorzeichen unterscheiden. Ihr Betrag kann aus der Aufspaltungsweite entnommen werden. Man findet $\mu_z = \pm e\hbar/2m \equiv \pm \mu_B$, wobei $\hbar$ die Plancksche Konstante bedeutet (↑ 1.2.3.). Dieser Wert μ_B wird als Bohrsches Magneton bezeichnet. Da μ und s proportional zueinander sind, hat auch die Feldkomponente des Spins nur zwei mögliche Werte. Aus der Quantenmechanik folgt $s_z = \pm \hbar/2$.

1.1.3. Elektronen- und Ionenstrahlen

1.1.3.1. Elementarprozesse zur Erzeugung freier Ladungsträger. Die Erzeugung von Elektronen an Festkörperoberflächen kann durch thermische Emission, Feldemission, Fotoemission und Sekundärelektronenemission erfolgen. Bei der *thermischen Emission* wird den quasifreien Elektronen des Festkörpers durch Erhitzen desselben die zur Überwindung des Potentialwalls an der Festkörperoberfläche notwendige materialspezifische *Austrittsenergie* zugeführt. Dieser Wert vermindert sich bei Anliegen eines elektrischen Feldes E_0 an dem als Katode geschalteten festen Körper um den Betrag $e\sqrt{eE_0/4\pi\varepsilon_0}$. Die *Feldemission* von Elektronen setzt ein, wenn die Feldstärke an der Festkörperoberfläche einige 10^9 Vm^{-1} erreicht. In diesem Falle durchtunneln die quasifreien Elektronen den Potentialwall an der Festkörperoberfläche, dessen Dicke nunmehr mit der Materiewellenlänge (↑ 1.2.3.) der Elektronen vergleichbar ist, ohne Energieaufnahme. Bei der *Fotoemission* wird dem quasifreien Elektron des Festkörpers die für den Austritt aus der Oberfläche notwendige Energie von einem absorbierten Photon übertragen. Die *Sekundärelektronenemission* wird durch die Einwirkung von Primärelektronen bzw. Atomen und Ionen genügend hoher Energie auf die Festkörperoberfläche hervorgerufen.
Positive Ionen werden durch Herauslösen eines oder mehrerer Elektronen aus der Elektronenhülle neutraler Atome oder Moleküle erzeugt. Die erforderliche Ionisierungsenergie E_i wird den neutralen Atomen oder Molekülen durch unelastische Wechselwirkung mit energetischen Teilchen übertragen. Die dafür notwendige Bedingung erhält man aus dem Energie- und Impulserhaltungssatz:

$$\Delta E_{\max} = \frac{m_2}{m_1 + m_2} E_1 \geqq E_i,$$

wobei m_2 die Masse des neutralen Atoms oder Moleküls, m_1 die Masse und E_1 die Energie des stoßenden Teilchens und $\Delta E_{\max}$ die maximal übertragbare Energie ist. Die effektivste Methode zur Erzeugung von Ionen aus Gasen ist die *Elektronenstoßionisation*, da wegen $m_e = m_1 \ll m_2$ das Elektron seine kinetische Energie beim Stoß fast vollständig auf das Atom oder Molekül übertragen kann. Die *Ionenstoßionisation* ist wegen $m_i = m_1 \approx m_2$ weniger effektiv. Die *Fotoionisation* von Atomen und Molekülen findet statt, wenn die Energie $h\nu$ des absorbierten Photons mindestens der

Ionisierungsenergie E_i entspricht. Bei genügend hohen Temperaturen in Gasen erhalten einige neutrale Atome und Moleküle eine ausreichende Energie, so daß der Ionisationsprozeß durch *Neutralteilchenstoß* einsetzt; die dabei abgelösten Elektronen, die wegen des thermischen Gleichgewichts die gleiche Energie annehmen, ionisieren anschließend mit wesentlich größerer Wahrscheinlichkeit als die Neutralteilchen. Diesen Vorgang bezeichnet man als *thermische Ionisation*. Die Erzeugung positiver Ionen an festen Grenzflächen kann durch *Sekundärionenemission* beim Beschuß fester Körper mit energetischen Ionen (Zerstäubungsprozeß), durch *Feldionisation* von Atomen und Molekülen an festen Grenzflächen, durch *Ionisation an heißen Oberflächen* und durch *thermisches Verdampfen* erfolgen. Die Bildung negativer Ionen gelingt durch Anlagerung langsamer Elektronen oder negativer Ionen an neutrale Atome oder Moleküle.

1.1.3.2. Elektronenstrahlerzeugung. Als elektronenauslösende Prozesse zur Erzeugung von Elektronenstrahlen werden vorwiegend die thermische Emission und die Feldemission angewendet. Die emittierten Elektronen werden durch ein Beschleunigungsfeld zwischen der Katode und der gegenüberliegenden Anode zu einem Elektronenstrahl formiert (Abb. 1.1.3-1). Die Stromdichte und der Querschnitt des Elektronenstrahls können mit Hilfe einer zwischen Katode und Anode angeordneten Steuerelektrode, die ein veränderliches Potential gegenüber der Katode erhält, reguliert werden.

1.1.3.3. Ionenstrahlerzeugung. Die Ionen werden in einer separaten Ionisierungskammer erzeugt, danach extrahiert, beschleunigt und zu einem *Ionenstrahl* formiert. Dazu dient ein *Extraktionselektrodensystem*, das gewöhnlich aus Emissions- und Beschleunigungselektrode besteht. Ionenstromstärken im Mikroamperebereich, bespielsweise zur massenspektroskopischen Materialanalyse, werden u. a. durch *Elektronenstoßionenquellen*, *Ionenstoßionenquellen*, *Feldionisationsionenquellen* (insbesondere für kurzlebige Ionen), *Elektronenanlagerungsionenquellen* und verschiedene Arten von *Plasmaionenquellen* erzeugt. Die Ionenerzeugung in Gasentladungsplasmen (↑ 9.1.6.4.) erfolgt vorwiegend durch Elektronenstoßionisation. Ionenstrahlen mit großem Strahlquerschnitt und Ionenströmen bis zu einigen Ampere können mit speziell konstruierten Ionenquellen erzeugt werden. Bei einer *Duoplasmatronionenquelle* (Abb. 1.1.3-2) brennt eine Bogenentladung zwischen der Glühkatode und der gekühlten Anode durch den Kanal der Zwischenelektrode hindurch. Die Zwischenelektrode und die Anode sind zur Erzeugung eines starken inhomogenen Magnetfelds als Polschuhe ausgebildet. Der dadurch hochverdichtete Teil der Entladung bildet das Anodenplasma mit Teilchendichten bis 10^{22} m^{-3}. Die Anode stellt gleichzeitig die Emissionselektrode dar.

Bei einer *Niederdruckionenquelle* (Finkelsteintyp) mit magnetisch geführter Elektronenpendelung (Abb. 1.1.3-3) besteht das Elektrodensystem aus der Glühkatode, dem Anodenzylinder und der Reflexionselektrode, die gleichzeitig als Emissionselektrode ausgebildet ist. Durch die Wirkung des elektrischen und magnetischen Felds pendeln die von der Glühkatode emittierten Primärelektronen in einer Potentialmulde auf Spiralbahnen zwischen Glüh- und Refle-

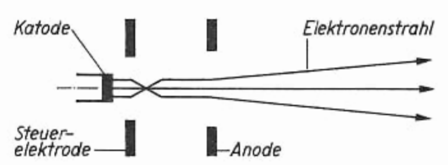


Abb. 1.1.3-1 Elektrostatischer Elektronenstrahlerzeuger

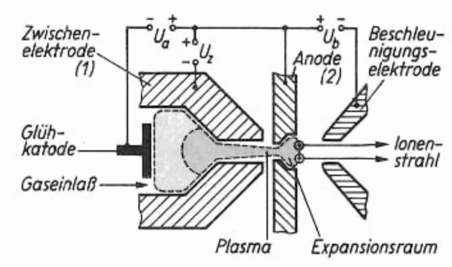


Abb. 1.1.3-2 Elektrodenanordnung einer Duoplasmatronionenquelle; U_a Anodenspannung, U_z Spannung an der Zwischenelektrode, U_b Beschleunigungsspannung, (*1*) und (*2*) stellen zugleich Teile einer Polschuhlinse dar

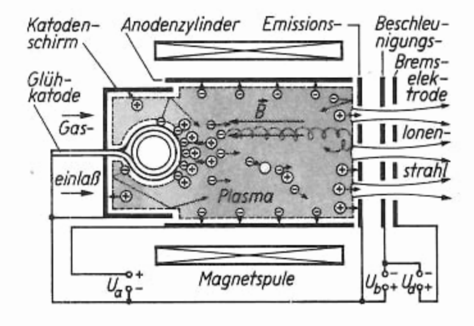


Abb. 1.1.3-3 Elektrodenanordnung und Wirkprinzip einer Finkelsteinionenquelle; U_a Entladungsspannung, U_b Beschleunigungsspannung, U_d Bremsspannung

xionselektrode, wodurch sich ihre Weglänge und damit die Ionisierungswahrscheinlichkeit erheblich vergrößern. Die Ionenextraktion erfolgt hier mit Hilfe eines Dreielektroden-extraktionssystems. Die Bremselektrode dient zur Steuerung der Ionenenergie.

1.1.3.4. Elektronen- und Ionenstrahlführung. Alle Anwendungen von Elektronen- und Ionen-

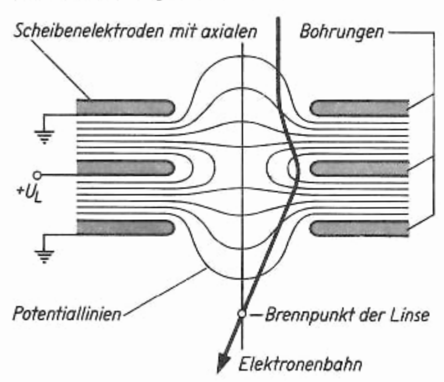
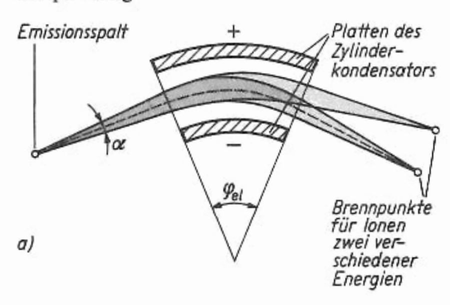


Abb. 1.1.3-4 Schematische Darstellung des Potentialfeldverlaufs und der Elektronenbahn in einer elektrostatischen Einzellinse; U_L Linsenspannung



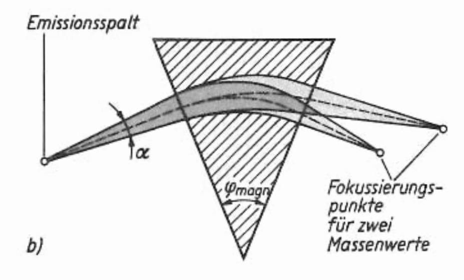


Abb. 1.1.3-6 a) Geschwindigkeitsfokussierung im elektrischen Sektorfeld, b) Richtungsfokussierung und Massentrennung im magnetischen Sektorfeld; α Öffnungswinkel des Ionenstrahls, φ Feldwinkel

strahlen haben gemeinsam, daß geladene Teilchen über eine vorgegebene Strecke durch das Vakuum geführt werden müssen. Das Strahlführungssystem hat die Aufgabe, die Parameter des in der Quelle formierten Elektronen- bzw. Ionenstrahls in die erforderlichen Strahlparameter am Prozeßort zu transformieren. Dazu werden elektrische und magnetische Beschleunigungs-, Ablenk-, Fokussierungs- und Selektierungssysteme verwendet. Elektromagnetische Felder wirken dabei auf die Ladungsträgerstrahlen ähnlich wie optische Linsen, Prismen und Spiegel auf Lichtstrahlen. Die Beschleunigung der Ladungsträgerstrahlen im elektromagnetischen Feld folgt aus der Lorentzkraft (↑ 2.1.6.5.); für die Ablenkung im elektrischen Feld ist die Teilchenenergie und für die Ablenkung im Magnetfeld der Impuls der Teilchen maßgebend (↑ 2.1.6.6.). Zur Abbildung und Fokussierung von Ladungsträgerstrahlen werden vorwiegend rotationssymmetrische elektromagnetische Felder verwendet. Der aus den Gesetzen der Ladungsträgerbewegung folgende *Hauptsatz der Ladungsträgeroptik* lautet: Bündel von flachen und achsennahen Ladungsträgerbahnen, die von den Punkten eines achsensenkrechten und auf die Umgebung der Achse beschränkten Ebenenstücks ausgehen, vereinigen sich nach dem Durchgang durch ein beliebiges rotationssymmetrisches elektromagnetisches Feld wieder in einem Punkt. Die Vereinigungspunkte bilden wieder ein achsensenkrechtes Ebenenstück, das ein ähnliches Abbild der Ausgangsfläche darstellt.

Abbildende und fokussierende rotationssymmetrische elektrische oder magnetische Felder werden mit ladungsträgeroptischen Linsen erzeugt. Die einfachste *elektrostatische Linse* besteht aus drei Scheibenelektroden mit zentralen Bohrungen auf einer Achse; durch unterschiedliche elektrische Potentiale an den Elektroden bildet sich ein elektrisches Feld mit fokussierender Wirkung aus (Abb. 1.1.3-4). Die einfachste *magnetische Linse* ist eine kurze stromdurchflossene zylindrische Spule (Abb. 1.1.3-5). Ihre Brechkraft kann durch eine Eisenkapselung, die als Polschuhpaar ausgebildet ist, erhöht werden. Selektierungssysteme werden durch elektrostatische und magnetische Sektorfelder gebildet (Abb. 1.1.3-6).

Die bisherigen Ausführungen berücksichtigen nicht die Raumladungswirkung der Ladungsträger. Elektronen- und Ionenstrahlen bilden durch die Wirkung ihrer eigenen Raumladung ein divergentes Strahlenbündel. Durch geeignete geometrische Elektrodenanordnungen läßt sich diese Divergenz im Strahlerzeuger unterdrücken (Piercestrahler). Eine wenigstens teilweise Kompensation der Raumladung können Ladungsträger des umgekehrten Vorzeichens bewirken, die beispielsweise durch Restgasioni-

sation oder in separaten Teilchenquellen erzeugt werden.

1.1.3.5. Elektronen- und Ionenstrahlgeräte. Für die Anwendung in Wissenschaft und Technik ist eine große Zahl von Elektronen- und Ionenstrahlgeräten entwickelt worden (Tab. 1.1.3-7 und 1.1.3-8). Einige Gerätegruppen werden nachfolgend kurz beschrieben:

Die *Elektronenstrahlröhre* (Abb. 1.1.3-9) dient zur bildlichen Darstellung elektrischer Signale. Sie besteht im wesentlichen aus dem Elektronenstrahlerzeuger, dem Ablenksystem und dem Leuchtschirm. Die aus der Katode austretenden Elektronen werden mit Hilfe elektrostatischer oder magnetischer Linsen auf dem fluoreszierenden Leuchtschirm als sichtbarer scharfer Fleck abgebildet. Eine vor der Katode angebrachte Steuerelektrode ermöglicht die Intensitätssteuerung des Elektronenstrahls. Durch horizontal und vertikal gerichtete elektrische oder magnetische Querfelder, deren Feldstärken den darzustellenden Signalen proportional sind, läßt sich der Elektronenstrahl und somit der Leuchtfleck praktisch trägheitslos ablenken. Weiterentwicklungen der Elektronenstrahlröhren sind beispielsweise die Fernsehbildwiedergaberöhren.

Die *Elektronenmikroskope* ermöglichen die bildliche Darstellung der kleinsten Bereiche der Materie bis hin zu atomaren Größenordnungen (Tafeln 26/27). Die Elektronenstrahlführung und -fokussierung erfolgt mit Hilfe elektrischer oder magnetischer Linsen. In Abb. 1.1.3-10 ist der jeweilige Strahlengang einiger Elektronenmikroskope schematisch dargestellt. Das hohe

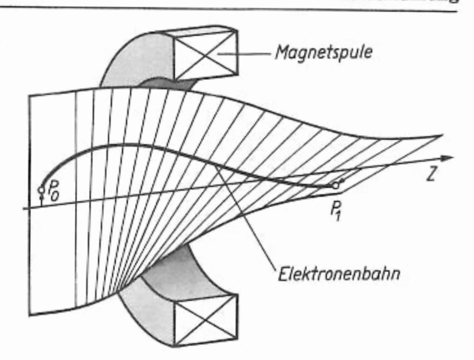


Abb. 1.1.3-5 Darstellung der Elektronenbahn durch ein rotationssymmetrisches Magnetfeld; von P_0 ausgehende Elektronenbahnen liegen auf einer Schraubenfläche mit veränderlicher Ganghöhe, Z Symmetrieachse, P_0 und P_1 Bahnpunkte

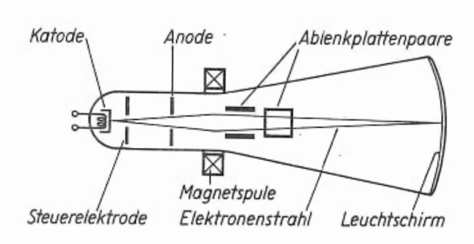


Abb. 1.1.3-9 Elektronenstrahlröhre

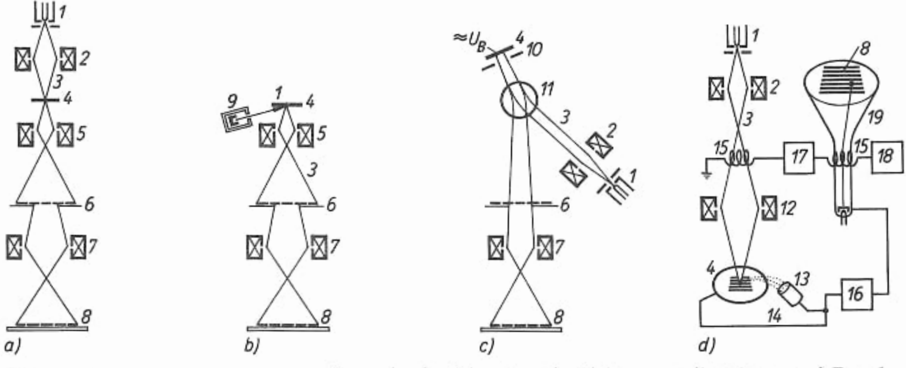


Abb. 1.1.3-10 Schematische Darstellung der Strahlengänge in Elektronenmikroskopen; a) Durchstrahlungselektronenmikroskop, b) Emissionselektronenmikroskop, c) Elektronenspiegelmikroskop, d) Rasterelektronenmikroskop; *1* Elektronenquelle, *2* Kondensor, *3* Elektronenstrahl, *4* Objekt, *5* Objektiv, *6* Zwischenbild, *7* Projektiv, *8* Endbild, *9* Ionenquelle (Elektronenquelle), *10* Elektronenspiegel, *11* magnetisches Umlenkfeld, *12* magnetische Linse, *13* Detektor, *14* Probenstrom, *15* Ablenkeinheit, *16* Videoverstärker, *17* Variation der Bildgröße, *18* Rastergenerator, *19* Bildröhre, U_B Beschleunigungsspannung des Elektronenstrahls

Tab. 1.1.3-8 Anwendung von Ionenstrahlen

Anwendungsgebiet	Geräteart oder Verfahren	Merkmale und Anwendungsbeispiele
Materialanalyse	Massenspektroskopische Apparate:	
	Massenspektrometer Massenspektrographen	Zerlegung von Ionenstrahlen nach q/m mit Hilfe elektrischer und magnetischer Felder
	Massentrenner	Zerlegung technisch verwertbarer Substanzmengen
	Sekundärionenmassenspektrometrie (SIMS)	Analyse der beim Ionenbeschuß von Festkörperoberflächen abgelösten Sekundärionen; Untersuchung der Zusammensetzung fester Körper und dünner Schichten
	Ionenrückstreuanalyse	Energieanalyse der beim Beschuß von Festkörperoberflächen rückgestreuten Primärionen; Untersuchung der Oberfläche
Materialumwandlung	Beschleuniger: Zyklotron Synchrotron	Erzeugung hochenergetischer Ionenstrahlen zur Untersuchung von Atomkern- und Elementarteilchenumwandlungsprozessen
Ionenabbildung	Feldionenemissionsmikroskop	Feldionisation von neutralen Gasteilchen an einer als Spitze ausgebildeten Festkörperoberfläche; Verwendung der entstehenden Ionen zur vergrößerten Darstellung der Festkörperoberfläche auf einem Leuchtschirm
Ionenstrahltechnologie	Ionenstrahlzerstäubungsablagerung	Beschichten von Substraten mit abgestäubtem Festkörpermaterial im Hochvakuum
	Ionenstrahlablagerung	Beschichten von Substraten durch Ablagerung ionisierter Teilchenströme im Hochvakuum
	Ionenstrahlplattieren	Beschichten von Substraten durch Ablagerung teilionisierter Teilchenströme
	Ionenstrahlinduzierte Schichtabscheidung	Modifizierung der Eigenschaften aufwachsender Schichten durch Ionenbeschuß
	Ionenimplantation	Einlagerung von Ionen in den oberflächennahen Bereich fester Körper; beispielsweise Dotierung von Halbleitern
	Ionenstrahlätzen	Abtragen von Festkörperoberflächen durch Ionenbeschuß; Strukturätzen in der Mikroelektronik
Energieerzeugung	Tokamak	Einschuß stromstarker Ionenstrahlen in eine Magnetfalle zur Untersuchung der gesteuerten Kernfusion
Raumfahrt	Ionentriebwerk	Ausnutzung des Impulses der auf hohe Energien beschleunigten Ionenstrahlen zum Antrieb von Raumflugkörpern außerhalb der Erdatmosphäre

Tab. 1.1.3-7 Anwendung von Elektronenstrahlen

Anwendungsgebiet	Geräteart oder Verfahren	Merkmale und Anwendungsbeispiele
Signalumwandlung und Signalverstärkung	Elektronenröhren:	
	Gleichrichterröhre	Nutzung des von der Katode zur Anode fließenden Elektronenstroms zur Gleichrichtung von Wechselströmen
	Verstärkerröhre, Senderöhre, Klystron, Magnetron, Wanderfeldröhre	Modellierung der Intensität von Elektronenströmen zur Signalverstärkung
	Fotozelle	Nutzung der Fotoelektronenemission an Festkörperoberflächen; Umwandlung von optischen Signalen in elektrische
	Sekundärelektronenvervielfacher (SEV)	Nutzung der Sekundärelektronenemission an Festkörperoberflächen; Registrierung geringer Teilchenströme
Material- und Strukturanalyse	Elektronenbeugungsgeräte zur:	
	Reflexionselektronenbeugung (REB) Durchstrahlungselektronenbeugung (TEB)	Ausnutzung der Wellennatur von Elektronenstrahlen zur Untersuchung der Struktur und des kristallinen Aufbaus von dünnen Schichten und festen Körpern
	Röntgengeräte	Untersuchung der Struktur von festen Körpern und Flüssigkeiten mit Hilfe von Röntgenstrahlen, die durch Elektronenbeschuß von Festkörperoberflächen erzeugt werden
	Elektronenstrahlmikroanalysator	Analyse der durch Elektronenbeschuß angeregten Röntgenstrahlung zur Materialidentifikation des festen Körpers
	Elektronenenergieverlustanalyse	Energieverlustanalyse der an einer Oberfläche reflektierten Elektronen zur Material- und Strukturaufklärung des festen Körpers
Elektronenabbildung	Elektronendurchstrahlungsmikroskop	Nutzung der durch eine dünne Schicht hindurchtretenden Elektronen eines Elektronenstrahls zur Abbildung derselben auf einen Leuchtschirm oder einer Fotoplatte
	Elektronenrückstrahlungsmikroskop	Nutzung der von der Oberfläche rückgestreuten Elektronen eines Elektronenstrahls zur Abbildung
	Elektronenspiegelmikroskop	Nutzung der vor einer Oberfläche reflektierten Elektronen eines Elektronenstrahls zur Darstellung des Ladungsprofils der Oberfläche
	Elektronenemissionsmikroskop Elektronenfeldemissionsmikroskop	Nutzung der von einer Festkörperoberfläche emittierten Elektronen zur Darstellung des Emissionsvermögens, der Oberflächenstruktur und der Oberflächenabbildung

	Elektronenrastermikroskop (Scanning-elektronenmikroskop)	Zeilenweise Abtastung einer Festkörperoberfläche mit einem Elektronenstrahl und Verwendung der emittierten Sekundärelektronen, der reflektierten Primärelektronen oder der erzeugten Röntgenstrahlung zur Abbildung der Oberfläche auf einem synchron mit dem Elektronenstrahl betriebenen Leuchtschirm
Bilddarstellung und Bildübertragung	Oszillographenröhren	Bildliche Darstellung elektrischer Signale auf einem Leuchtschirm
	Fernsehkameraröhren	Umwandlung eines Bildes in elektrische Signale
	Fernsehbildwiedergaberöhren	Umwandlung elektrischer Signale in ein Bild auf einem Leuchtschirm
Materialumwandlung Strahlenchemie	Teilchenbeschleuniger:	
	Van-de-Graff-Generator	Erzeugung hochenergetischer Elektronenstrahlen zur Untersuchung der Wechselwirkung von Elektronen mit dem Atomkern und zur Erzeugung energiereicher γ-Quanten
	Betatron	Strahlenchemische Anwendungen: Vernetzen, Vulkanisieren, Polymerisieren, Sterilisieren
	Elektronenzyklotron	
	Elektronensynchrotron	
Elektronenstrahltechnologie	Elektronenstrahlverdampfen	Hochratebeschichtung von Festkörperoberflächen mit kontaminationsfreien Schichten
	Elektronenstrahlschmelzen	Herstellung ultrareiner Legierungen; Schmelzen von Materialien mit hohem Schmelzpunkt
	Elektronenstrahlschweißen	Erzeugung hochqualitativer schmaler Schweißnähte
	Elektronenstrahlbearbeiten:	
	Mikrostrukturieren	Technologische Nutzung der thermischen und nichtthermischen Wirkung von Elektronenstrahlen zur Mikrobearbeitung von Werkstücken und Schichten
	Schneiden	
	Bohren	
	Bestrahlen	
	Aufheizen	
	Härten	

Auflösungsvermögen dieser Geräte beruht auf der um Größenordnungen kleineren Wellenlänge der Elektronen (↑ 1.2.3.3.) gegenüber Licht, da das Auflösungsvermögen durch mit der Wellenlänge anwachsende Beugungseffekte (↑ 2.3.2.) begrenzt wird. Zwei Objektpunkte sind noch gut zu unterscheiden, wenn das Maximum des Beugungsbilds des einen mit dem Minimum des anderen zusammenfällt. Das *Feldemissionsionenmikroskop* ermöglicht die Darstellung des Oberflächenbildes von Festkörperspitzen; es besitzt ein Punktauslösungsvermögen von etwa 0,2 nm, so daß die Lage einzelner Atome sichtbar gemacht werden kann (Tafel 30).
Massenspektroskopische Apparate dienen zur Analyse von ionisierten Teilchenströmen und zur Massentrennung. Sie spalten einen Ionenstrahl mit Hilfe elektrischer und magnetischer Felder nach den verschiedenen q/m_i-Werten der Teilchen zum Massenspektrum auf. Im *Massenspektrometer* erfolgt die Registrierung der Massenkomponenten des Ionenstrahls nacheinander mit elektrischen Methoden; im *Massenspektrographen* werden die Massenkomponenten gleichzeitig auf einer Fotoplatte registriert. In doppelfokussierenden Apparaten (Abb. 1.1.3-11) werden ein elektrisches Radialfeld und ein magnetisches Sektorfeld so miteinander kombiniert, daß Ionen gleicher Masse und Ladung, jedoch mit verschiedener Anfangsenergie und Einschußrichtung in einem Fokus der Registrierebene wieder vereint werden. Das als Verhältnis von mittlerer Ionenmasse zu

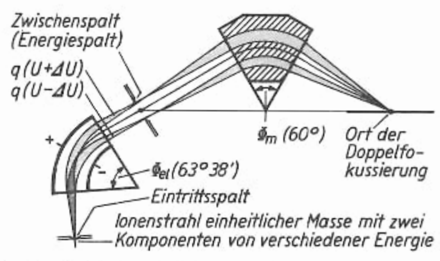


Abb. 1.1.3-11 Doppelfokussierendes Massenspektrometer; Φ_{el} Winkel des elektrischen Felds, Φ_m Winkel des magnetischen Felds, qU mittlere Ionenenergie

noch nachweisbarer Massendifferenz definierte *Auflösungsvermögen* dieser Geräte beträgt $(1 \dots 5) \cdot 10^5$. Neben der Massenspektroskopie positiver Ionen, die vorwiegend durch Elektronenstoßionisation erzeugt werden, ist die *Elektronenanlagerungsmassenspektroskopie* entwickelt worden. Dieses Verfahren eignet sich insbesondere zur Bestimmung der Molekulargewichte organischer Substanzen, da eine Dissoziation durch Elektronenstoß vermieden wird.

Apparate zur Analyse stromstarker Ionenstrahlen technisch verwertbarer Substanzmengen werden als *Massentrenner* bezeichnet. Zur Dotierung von Halbleitermaterialien (↑ 8.3.3.1.) verwendet man in zunehmendem Maße massenanalysierte Ionenstrahlen aus den Dotanden, die in *Ionenimplantationsanlagen* erzeugt, auf hohe Energie beschleunigt und in den Halbleiter implantiert werden. Für die Herstellung dünner Schichten und zur Mikrobearbeitung von Festkörperoberflächen gewinnen die Verfahren der *Ionenstrahltechnologie* wachsende Bedeutung (Tafeln 22/23 und 24/25).

1.1.4. Strukturiertes Atom

1.1.4.1. Elektronen und Atomkern als Atombausteine. Gegen Ende des 19. Jahrhunderts bestand unter den Physikern allgemein Klarheit darüber, daß Atome weiter zerlegbar sind und eine komplizierte Struktur unter Beteiligung des Elektrons besitzen. Solche Ansichten wurden vor allem durch eine Fülle experimenteller Ergebnisse zu den Atomspektren gestützt. Eine Reihe namhafter Physiker wie Stoney, Lorentz und Perrin hatten konkretere, am Bau des Sonnensystems orientierte Vorstellungen von der Struktur des Atoms. Nach dem 1904 von Thomson formulierten Atommodell dagegen sollten sich die Elektronen vom Durchmesser 10^{-15} m im Inneren einer homogen positiv geladenen Kugel der Ausdehnung 10^{-10} m befinden und durch ihre Schwingungen die Emission von Licht verursachen. Aber schon Lenard hatte aus seinen Experimenten mit Katodenstrahlen gefolgert, daß das „Innere" des Atoms im wesentlichen „leer" sein müsse. Die endgültige Entscheidung zugunsten des Planetenmodells führte 1911 Rutherford herbei.
Rutherford, Geiger und Marsden untersuchten den Durchgang von α-Teilchen, d. h. zweifach positiv geladenen Heliumkernen, durch Metallfolien. Aus der von einer radioaktiven Quelle isotrop emittierten α-Strahlung wurde mit Hilfe von Blenden ein schmales Bündel entnommen. Die Auftreffpunkte der einzelnen α-Teilchen, die die Metallfolie durchdrungen hatten, konnten als Lichtblitze auf einem Zinksulfidschirm unter dem Mikroskop ausgezählt werden. Dieses Nachweisverfahren, das Crookes 1903 entwickelt hatte, ist unter dem Namen *Szintillationsmethode* bekannt geworden. Im Experiment wurde teilweise eine starke Ablenkung der α-Teilchen beobachtet, ja sogar vereinzelt Rückstreuung. Rutherford zeigte rechnerisch, daß solche Ablenkungen nur entstehen können, wenn im Gegensatz zum Thomsonschen Modell ein massereiches Zentrum bis zu Abständen r der Größenordnung

10^{-14} m herab ein elektrostatisches Feld proportional zu $1/r^2$ erzeugt.
Als einzig mögliche Interpretation dieser Ergebnisse formulierte RUTHERFORD 1911 in einem Vortrag die Prinzipien des heute nach ihm benannten Atommodells: Das Atom besteht aus einem Z-fach positiv geladenen Kern vom Durchmesser 10^{-14} bis 10^{-15} m, in dem fast die gesamte Atommasse vereinigt ist, und Z Elektronen, die ihn etwa im Abstand 10^{-10} m umkreisen und die Atomhülle bilden. Das *Rutherfordmodell* fand zunächst wenig Anhänger, denn nach den Gesetzen der Elektrodynamik müßten die mit einer Radialbeschleunigung von 10^{23} ms^{-2} umlaufenden Elektronen nach etwa 10^{-8} s ihre Energie in Form elektromagnetischer Wellen mit kontinuierlichem Spektrum abgestrahlt haben und in den Kern gestürzt sein. Zwar strahlen die Atome nach vorheriger Anregung tatsächlich, die spektrale Verteilung der emittierten Strahlung zeigt aber charakteristische Linienspektren, und schließlich wird ein stabiler, nicht mehr strahlender Zustand des Atoms erreicht.

1.1.4.2. Entwicklung der Atomspektroskopie. Mindestens seit dem 18. Jahrhundert ist bekannt, daß weißes Licht beim Durchgang durch ein Glasprisma in seine farbigen Bestandteile zerlegt wird. Der Ausbau dieser Methode zu einem der empfindlichsten und vielseitigsten physikalischen Meßverfahren begann 1814 mit der Untersuchung des Sonnenspektrums durch FRAUNHOFER. Er fand 567 der nach ihm benannten dunklen Linien im ansonsten kontinuierlichen Spektrum des Sonnenlichts. Die markantesten Linien bezeichnete er durch große lateinische Buchstaben. Die besonders auffällige *doppelte D-Linie* im gelben Teil des Spektrums wurde zum Ausgangspunkt einer Reihe wichtiger Entdeckungen. KIRCHHOFF und BUNSEN verglichen das Sonnenspektrum mit dem Spektrum einer Spiritusflamme. Es zeigte sich, daß im Spektrum der Flamme eine gelbe Doppellinie aufleuchtet, wo im Sonnenspektrum die dunkle doppelte D-Linie liegt. Statt des Sonnenlichts untersuchten sie nun das von einem glühenden Körper ausgestrahlte Licht, das ein kontinuierliches Spektrum besitzt, welches nicht von dunklen Linien unterbrochen wird. Die dunkle D-Linie trat aber sofort wieder auf, als sie das Glühlicht vor seiner spektralen Zerlegung durch die Spiritusflamme treten ließen. Also absorbiert die Flamme Licht von genau derselben Wellenlänge als sie emittiert.
Ein weiteres wichtiges Ergebnis der Arbeiten von KIRCHHOFF und BUNSEN war, daß die gelbe Doppellinie bzw. die dunkle Fraunhofersche D-Linie mit Sicherheit auf die Beteiligung des Elements Natrium bei der spektralen Gestaltung des untersuchten Lichts schließen lassen.
Die gewonnenen Erkenntnisse verallgemeinerte KIRCHHOFF 1859 zu folgenden Aussagen: Jedes Element hat sein Linienspektrum mit charakteristischer Linienfolge. Ausgehend von dieser Feststellung entwickelte sich die *Spektralanalyse*, die auch bei geringen Substanzmengen höchste Genauigkeit erreicht. Eine Reihe von Elementen wurde auf spektralanalytischem Wege entdeckt, z. B. Rubidium und Caesium durch KIRCHHOFF und BUNSEN, Gallium durch DE BOISBAUDRAN und 1868 Helium durch JANSSEN in der Sonnenkorona sowie 1895 durch RAMSAY auf der Erde. Die Spektralanalyse wurde zur weitaus leistungsfähigsten Untersuchungsmethode in der Atom- und Astrophysik.
Die Sonne besteht aus einem heißen Kern, der Licht mit kontinuierlichem Spektrum emittiert, und einer kühlen Gashülle, die durch Absorption die Fraunhoferschen Linien erzeugt. Natrium ist Bestandteil der Sonnenkorona.
Diese These wurde 1868 durch Beobachtungen des französischen Astronomen JANSSEN bestätigt, die er anläßlich einer totalen Sonnenfinsternis machte. Im Licht der Sonnenkorona fand er die D-Linie des Natriums als hell leuchtende Doppellinie. Also verhält sich die Sonnenkorona wie die Spiritusflamme. In der Folgezeit wurden die Spektren aller bekannten Atomsorten katalogisiert. Dabei verbesserte sich das Auflösungsvermögen der Spektralapparate wesentlich durch Einführung von Beugungsgittern anstelle der Glasprismen. Außerdem konnten nun Bereiche des Spektrums erschlossen werden, für die Glasprismen wegen ihrer starken Absorption ungeeignet sind.
Die Atom- und Molekülspektroskopie ist ständig weiterentwickelt worden und stellt heute ein extrem leistungsfähiges Verfahren dar, mit dem prinzipiell einzelne Atome nachgewiesen und weitreichende Schlußfolgerungen auf die physikalischen Parameter in der Umgebung der strahlenden Atome gezogen werden können. Letzteres ist von entscheidender Bedeutung für die Astrophysik, die fast ausschließlich auf die Analyse der elektromagnetischen Strahlung extraterrestrischer Objekte als Informationsquelle angewiesen ist. Für die Entwicklung bedeutsam waren einmal die Vervollkommnung der experimentellen Methoden, gekennzeichnet durch die Erhöhung des spektralen Auflösungsvermögens, die Erschließung völlig neuer Spektralbereiche (z. B. Mikrowellen) und den Einsatz von Lasern als Strahlungsquellen großer Intensität bei sehr genau fixierter Frequenz. Zum anderen wurden die Umgebungseinflüsse systematisch erforscht, vor allem ausgehend von dem von ZEEMAN 1896 und STARK 1913 experimentell erbrachten Nachweis der Aufspaltung von Spektrallinien in magnetischen bzw. elektrischen Feldern. Schließlich schuf der

Aufbau der Quantenmechanik, beginnend mit der Bohrschen Theorie des Wasserstoffatoms (↑ 1.1.4.5.), die Voraussetzungen für die theoretische Analyse der Spektren und damit auch für die Auswertung immer feinerer Details, wie Aufspaltungen und Profil der Spektrallinien (Tafel 67).

1.1.4.3. Linienspektren und Niveaustruktur. Die schnell wachsende Menge experimentellen Materials zu den Atomspektren mußte wertvolle Hinweise auf die innere Struktur des Atoms enthalten. Es galt, die Gesetze der Linienverteilung aufzufinden. Einen wesentlichen Beitrag hierzu leistete 1885 BALMER, indem er

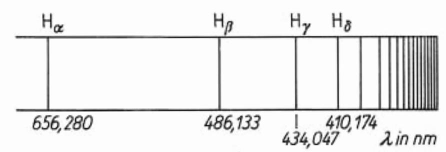


Abb. 1.1.4-1 Balmerserie des Wasserstoffspektrums

die Wellenlängen λ_n der Linienfolge im sichtbaren Teil des Wasserstoffspektrums durch die Formel $\lambda_n = \lambda_\infty n^2/(n^2 - 2^2)$ mit $\lambda_\infty = 364{,}56$ nm auf die ganzen Zahlen $n = 3, 4, 5, \dots$ zurückführte. Die unendlich vielen Linien H_α, H_β, $H_\gamma, \dots$ dieser als *Balmerserie* bezeichneten Linienfolge häufen sich an der Seriengrenze λ_∞ (Abb. 1.1.4-1). Wegen der endlichen Linienbreite und des begrenzten Auflösungsvermögens des benutzten Spektralapparats kann die Übereinstimmung der Balmerschen Formel mit dem Experiment nur bis zu einem maximalen n nachgewiesen werden. Damals erreichte man $n = 16$, heute dagegen $n = 35$.

1890 führte RYDBERG folgende Schreibweise der Balmerschen Formel ein: $1/\lambda_n = R(1/2^2 - 1/n^2)$, wobei $R = 109\,677{,}581\ \text{cm}^{-1}$ die *Rydbergkonstante* ist. Diese Darstellung verdeutlicht, daß jede Linie der Balmerserie durch Kombination des *Laufterms* $T_n = R/n^2$ mit dem *Festterm* $T_2 = R/2^2$ entsteht. RITZ verallgemeinerte 1908 diese Erkenntnis zum *Kombinationsprinzip*, das in seiner modernen Formulierung in der folgenden Feststellung besteht: Die Frequenz ν_{mn} einer Spektrallinie im Spektrum eines beliebigen Atoms ergibt sich durch Kombination zweier Terme entsprechend $\nu_{mn} = c(T_m - T_n)$. Mit Hilfe des Kombinationsprinzips konnte jeder Atomart ihr charakteristisches Termschema zugeordnet werden, so daß die Fülle der Spektrallinien aus vergleichsweise wenigen Termen folgt.

Besonders einfach war das Ergebnis der Analyse des Wasserstoffspektrums. Neben der Balmerserie mit dem Festtermindex $m = 2$, die sich vom Sichtbaren ins Ultraviolette erstreckt, wurden 1908 im Ultraroten die *Paschenserie* mit $m = 3$ und 1916 im fernen Ultraviolett die *Lymanserie* mit $m = 1$ entdeckt. BRACKETT und PFUND fanden zwei weitere Serien im Ultraroten mit Festtermen zu $m = 4$ und $m = 5$.

Ganz allgemein zeigen die Atomspektren je nach dem Auflösungsvermögen des Spektralapparats drei Strukturierungsebenen: *Grobstruktur*, *Feinstruktur* und *Hyperfeinstruktur*. Manche Linien, die bei geringer Auflösung als einheitlich erscheinen, zerfallen bei verbesserter Auflösung in eine Gruppe dicht benachbarter Linien, ein Multiplett, und lassen so ihre Feinstruktur erkennen. Bei weiterer Erhöhung der Auflösung findet man, daß diese Feinstrukturkomponenten aus mehreren Hyperfeinstrukturkomponenten bestehen können. Die D-Linie des Natriums ist ein typisches Beispiel. Sie besteht aus den Feinstrukturkomponenten D_1 und D_2 mit den Wellenlängen 589,592 4 nm und 588,995 0 nm. Jede der beiden Linien ist ein Hyperfeinstrukturdublett, dessen Komponenten sich nur um 0,0023 nm bzw. 0,0021 nm unterscheiden. Da nach dem Kombinationsprinzip Spektrum und Termschema eines Atoms eng miteinander zusammenhängen, muß auch das Termschema Grob-, Fein- und Hyperfeinstruktur besitzen. Für die typischen Termabstände der drei Strukturierungsebenen findet man $10^4\ \text{cm}^{-1}$, $10\ \text{cm}^{-1}$ und $10^{-1}\ \text{cm}^{-1}$.

1.1.4.4. Diskrete Atomzustände. FRANCK und G. HERTZ untersuchten im Jahre 1913 mit der in Abb. 1.1.4-2a dargestellten Versuchsanordnung Stoßprozesse zwischen den von der Glühkatode emittierten Elektronen und Quecksilberatomen im Grundzustand, also im Zustand kleinster innerer Energie. Der Druck des Quecksilberdampfs ist so bemessen, daß die Abstände Katode – Gitter *1* und Gitter *1* – Gitter *2* klein bzw. groß gegen die mittlere freie Weglänge der Elektronen sind. Stöße finden daher nur im feldfreien Raum zwischen Gitter *1* und Gitter *2* statt. Wegen der stark unterschiedlichen Massen verliert das Elektron nur dann Energie, wenn das Hg-Atom seine innere Energie auf Kosten der Bewegungsenergie des stoßenden Elektrons erhöht. Das

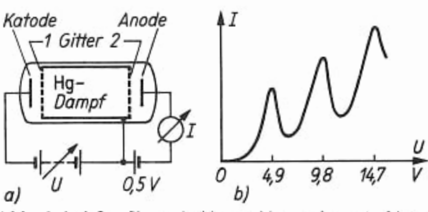


Abb. 1.1.4-2 Franck-Hertz-Versuch; a) Versuchsaufbau, b) Strom-Spannungs-Kennlinie

kleine Bremspotential zwischen Gitter *2* und Anode sorgt dafür, daß bei Erhöhung der Spannung U einsetzende unelastische Stöße als Rückgang des Stroms I bemerkt werden. Denn Elektronen, die alle ihre Bewegungsenergie an Hg-Atome abgegeben haben, können das Bremspotential nicht überwinden.
Im Experiment traten unelastische Stöße erstmals bei einer Elektronenenergie von 4,9 eV auf (Abb. 1.1.4-2b). Gleichzeitig emittierte der Hg-Dampf ein ultraviolettes Licht der Wellenlänge 253,65 nm. Die Minima bei den Elektronenenergien 9,8 eV, 14,7 eV, ... entstehen, da das Elektron genügend Energie besitzt, um nacheinander 2, 3, ... Hg-Atome anzuregen. Damit war gezeigt, daß Atome Energie nur in bestimmten Beträgen aufnehmen können. Jede Atomart besitzt daher ein für sie charakteristisches System diskreter zugelassener Zustände mit bestimmten inneren Energien. Der Übergang vom Grundzustand der Energie E_0 zum angeregten Zustand der Energie $E_n > E_0$ tritt nur ein, wenn dem Atom die Anregungsenergie $E_n - E_0$ zugeführt wird.

1.1.4.5. Bohrsche Bahnen und Quantensprünge. Die Grundlagen der Theorie der Atomhülle wurden 1913 von Bohr am Beispiel des aus einem Proton und einem Elektron bestehenden Wasserstoffatoms entwickelt. Nach der klassischen Physik ist die Bewegung des Elektrons auf einer Kreisbahn um das Proton durch das Gleichgewicht von anziehender Coulombkraft und abstoßender Zentrifugalkraft bestimmt. Bahnradius r und Bahngeschwindigkeit v hängen daher durch die Beziehung $e^2/4\pi\varepsilon_0 r^2 = mv^2/r$ miteinander zusammen. Dieses Gleichgewicht wird gestört, da das beschleunigt bewegte Elektron Energie in Form elektromagnetischer Strahlung abgibt und so seinen Bahnradius ständig verringert. Bohr beseitigte diese Schwierigkeit, indem er die klassische Physik innerhalb des Atoms durch zwei Postulate ergänzte. Im *1. Postulat* nimmt er noch vor der experimentellen Demonstration dieser Tatsache durch Franck und Hertz an: Es gibt im Atom sogenannte *stationäre Bahnen* um den Atomkern, auf denen sich Elektronen bewegen können, ohne zu strahlen. Diese stationären Bahnen wurden durch die folgende Quantenbedingung ausgewählt: Der Drehimpuls des Elektrons soll ein ganzzahliges Vielfaches der Planckschen Konstante $\hbar = h/2\pi$ betragen, also $mvr = n\hbar$ mit $n = 1, 2, 3, \ldots$ Damit findet man für die Radien der stationären Bahnen $r_n = n^2 a_0$, wobei $a_0 = \hbar^2 4\pi\varepsilon_0/e^2 m = 0{,}52917 \cdot 10^{-10}$ m als kleinstmöglicher Bahnradius die Bezeichnung Bohrscher Radius trägt. Er stellt eine für die Ausdehnung aller Atome charakteristische Länge dar. Die Energie des Elektrons im Wasserstoffatom ergibt sich als Summe von kinetischer Energie und potentieller Energie auf seinen stationären Bahnen zu $E_n = -(e^2/8\pi\varepsilon_0 a_0)(1/n^2)$.
Das *2. Postulat* besagt, daß der Übergang eines Elektrons von einer stationären Bahn der Energie E_n zu einer anderen der Energie E_m sprunghaft erfolgt, wobei die Energiedifferenz vom Atom als Lichtquant emittiert bzw. absorbiert wird, je nachdem ob E_m kleiner oder größer als E_n ist.
Diese Übergänge werden *Quantensprünge* genannt. Entsprechend der *Einsteinschen Lichtquantenhypothese* (↑ 1.2.2.) ist $h\nu$ die Energie eines Lichtquants der Frequenz ν. Für die Spektralfrequenz des Übergangs muß daher $\nu_{mn} = |E_m - E_n|/h$ gelten. Danach hängen die Spektralterme T_n durch die Beziehung $T_n = -E_n/ch$ mit den diskreten Energien E_n zusammen (s. a. 1.1.4.4.). Der Grob-, Fein- und Hyperfeinstruktur der Spektren (↑ 1.1.4.3.) entspricht eine analoge Struktur der atomaren Energieniveaus; die charakteristischen Energiemaßstäbe hierfür sind 1 eV, 10^{-3} eV und 10^{-5} eV. Diese Struktur wird durch die verschiedenen Anteile der elektromagnetischen Wechselwirkung der Elektronen mit dem Kern sowie untereinander bestimmt. Entscheidend für die Lagen der Grobstrukturniveaus ist die elektrostatische Wechselwirkung. Die Feinstruktur dieser Niveaus wird durch magnetische Wechselwirkung erzeugt, und die Wechselwirkung der Elektronen mit dem magnetischen Dipolfeld des Kerns bildet die Hauptursache der Hyperfeinstruktur.
In Übereinstimmung mit dem Ritzschen Kombinationsprinzip liefert das 2. Bohrsche Postulat die Frequenzen des vom Wasserstoffatom emittierten Lichts $\nu_{mn} = c(e^2/8\pi\varepsilon_0 a_0 hc)(1/m^2 - 1/n^2)$ als Differenzen zweier Termwerte. Für $m < n$ entstehen die Formeln aller Spektralserien des Wasserstoffs. Dabei wird die Rydbergkonstante, $R_\infty = e^2/8\pi\varepsilon_0 a_0 hc = 109\,737{,}303\ \mathrm{cm}^{-1}$, auf bekannte Naturkonstanten zurückgeführt. Diese Berechnung der Rydbergkonstante war eine der Hauptstützen der Bohrschen Theorie. Der angegebene Wert weicht vom Meßergebnis (↑ 1.1.4.3.) vor allem wegen der vernachlässigten Mitbewegung des Atomkerns ab. Zur Berücksichtigung dieses Effekts ist nach der Mechanik die Elektronenmasse m durch die reduzierte Masse $mM/(m + M)$ von Elektron und Proton zu ersetzen.
Bohr erklärte also die Stabilität der Atome nicht im Sinne einer Zurückführung auf bereits bekannte Gesetze. Die Größe seiner Leistung besteht darin, daß er die physikalische Theorie um die Erkenntnis der Existenz solcher stationären Zustände der Atome bereicherte. Der Beitrag von Bohr zur Entwicklung der Physik ist somit von gleicher Art wie der von Planck, dem Begründer der Quantenphysik. Planck löste 1900 das Problem der spektralen Zusam-

mensetzung der Hohlraumstrahlung (s. a. 1.2.5.1.) ebenfalls durch ein Postulat, das die Grenzen der damaligen Physik sprengte: Die Energie eines Systems, das harmonische Schwingungen der Frequenz ν ausführt, kann sich nur um ganzzahlige Vielfache von $h\nu$ ändern. Die in diesem Zusammenhang erstmals aufgetretene *Plancksche Konstante h* erkannte BOHR als universelle Naturkonstante und nutzte sie für die Auswahl der stationären Elektronenbahnen der Atome. Zur Deutung der Linienspektren der Atome wurde von BOHR die Einsteinsche Lichtquantenhypothese herangezogen. So konnte die Berechnung dieser Spektren auf die Bestimmung der atomaren Energieniveaus zurückgeführt werden.

1.1.4.6. Korrespondenzprinzip und Quantenbedingungen. Die stationären Bohrschen Bahnen wurden aus der Gesamtheit aller klassisch möglichen Bahnen (ohne Strahlungseffekte) durch die *Quantenbedingung* $l_n = n\hbar$ mit der *Quantenzahl* $n = 1, 2, 3, \ldots$ für den Elektronendrehimpuls ausgewählt. Allgemein erfordert die Anwendung der Bohrschen Gedanken auf mikroskopische Systeme stets die Formulierung gewisser Quantenbedingungen. Diese schränken die in der klassischen Physik kontinuierlichen Wertevorräte verschiedener physikalischer Größen auf diskrete Spektren ein und berücksichtigen so Besonderheiten atomarer Systeme. Eine wesentliche Hilfe für das Auffinden solcher Quantenbedingungen war das 1923 ebenfalls von BOHR aufgestellte *Korrespondenzprinzip*. Es regelt die gegenseitigen Beziehungen von klassischer Physik und Quantenphysik. Danach müssen die Aussagen der Quantenphysik über ein System für große Werte der Quantenzahlen seines Zustands in diejenigen Aussagen übergehen, die die klassische Physik über das System macht.

Die Wirkung des Korrespondenzprinzips wird deutlich an Hand der Bohrschen Theorie des Wasserstoffatoms. Klassisch muß die Frequenz der vom Elektron emittierten Strahlung durch seine Umlauffrequenz T_{klass}^{-1} oder Vielfache davon gegeben sein. Für ein Elektron auf der n-ten stationären Bahn erhält man $T_{\text{klass}}^{-1} = v_n/2\pi a_n = l_n/2\pi m a_n^2 = 2cR_\infty/n^3$. Die Aussagen der Quantentheorie des Wasserstoffatoms (↑ 1.1.4.5.) stimmen für beliebiges n keinesfalls hiermit überein. Für großes n liefern sie jedoch $\nu_{n-1,n} \approx 2cR_\infty/n^3 = T_{\text{klass}}^{-1}$, d. h., die Quantelung des Bahndrehimpulses ist mit dem Korrespondenzprinzip vereinbar.

Bei der Bestimmung der Linienspektren der Atome aufgrund ihrer Energieniveaus zeigte es sich, daß nicht alle denkbaren Kombinationen atomarer Niveaus tatsächlich beobachtete Linien liefern, sondern daß dabei bestimmte *Auswahlregeln* zu berücksichtigen sind. Besonders bei der Begründung dieser Auswahlregeln leistete das Korrespondenzprinzip eine wertvolle Hilfe.

SOMMERFELD entwickelte das Bohrsche Atommodell weiter, indem er elliptische Elektronenbahnen zuließ und die relativistische Geschwindigkeitsabhängigkeit der Elektronenmasse berücksichtigte. Die Anwendung dieser häufig als „*ältere Quantentheorie*" bezeichneten und durch Quantenbedingungen und Korrespondenzprinzip charakterisierten Bohr-Sommerfeldschen Vorstellungen vom Aufbau der Atome mit mehreren Elektronen führten in zahlreichen Fällen zum qualitativen Verständnis ihrer Spektren. Dagegen gelang es selbst beim Heliumatom, das nur zwei Elektronen besitzt, nicht, die Energie seines Grundzustands zu berechnen. Die ältere Quantentheorie ist ein Gemisch von klassischen Vorstellungen und zusätzlichen Quantenbedingungen, das keine konsistente Beschreibung von Mikroobjekten ermöglicht. Diesem Anspruch wurde erst die *moderne Quantentheorie* gerecht, die auf klassische Vorstellungen konsequent verzichtet; sie entstand um 1925 aus den Beiträgen von DE BROGLIE, HEISENBERG, SCHRÖDINGER und BORN.

1.1.4.7. Innere Elektronen und Röntgenspektren. 1895 entdeckte RÖNTGEN bei Experimenten mit Katodenstrahlen eine durchdringende, ionisierende Strahlung, die *X*- oder *Röntgenstrahlen*. Sie entstehen z. B., wenn schnelle Elektronen auf schwere Kerne treffen, stark abgebremst werden und den größten Teil ihrer Bewegungsenergie als kurzwellige elektromagnetische *Bremsstrahlung* mit kontinuierlichem Spektrum abgeben. Bei Elektronenenergien von mindestens einigen keV tritt zusätzlich eine als *charakteristische Strahlung* bezeichnete Komponente auf, die 1908 von BARKLA entdeckt wurde. Diese weist ein diskretes Spektrum mit nur wenigen Linien auf. In der Folge abnehmender Frequenzen unterscheidet man K-, L-, M- und N-Serien, deren Linienverteilungen charakteristisch für die von Elektronen getroffene Atomsorte sind. MOSELEY fand 1913, daß die Frequenz einer herausgegriffenen Linie entsprechend $\nu = \text{const.}\,(Z - \sigma)^2$ monoton mit der Ordnungszahl wächst und von der chemischen Bindung des Elements nicht beeinflußt wird; σ ist die sogenannte Abschirmkonstante.

Wegen ihrer hohen Frequenz und der Unabhängigkeit von den chemischen Eigenschaften muß die charakteristische Strahlung durch Quantensprünge von Elektronen nach kernnahen stationären Bahnen entstehen. Schon 1914, also lange vor der Entdeckung des Pauliprinzips (↑ 1.1.4.8.) ging KOSSEL davon aus, daß kernnahe Elektronenbahnen normalerweise kein weiteres Elektron aufnehmen können. Für die Entstehung der charakteristischen

Strahlung fand er folgende heute noch gültige Erklärung: Das einfallende Elektron schlägt ein Atomelektron aus einer inneren Bahn, und ein weiter außen kreisendes Elektron fällt in die entstehende Lücke, während das Atom die Differenzenergie als hochfrequente Röntgenstrahlung emittiert. Zum Beispiel können die Linienfrequenzen der K-Serie in Übereinstimmung mit dem *Moseleyschen Gesetz* durch die empirisch gefundene Serienformel $\nu_{1n} = cR(Z - 1)^2(1/1^2 - 1/n^2)$ dargestellt werden. Diese entspricht der Serienformel für die Lymanserie des Wasserstoffspektrums (↑ 1.1.4.3.). Folglich entsteht die K-Serie wie diese durch Quantensprünge eines Elektrons zur kernnächsten stationären Bahn im elektrostatischen Feld der Ladung $(Z - 1)\,e$. Die Abschirmkonstante ist $\sigma = 1$, d. h., eine Elementarladung der wirklichen Kernladung Ze wird durch ein anderes auf der kernnächsten Bahn umlaufendes Elektron abgeschirmt.

So bestätigte die Untersuchung der Röntgenspektren durch MOSELEY die 1911 durch VAN DEN BROEK geäußerte Vermutung, daß die Ordnungszahl eines Elements im Periodensystem identisch mit der Kernladungszahl seiner Atome ist.

1.1.4.8. Periodensystem und Pauliprinzip. Im Grundzustand besitzt das Atom die kleinstmögliche innere Energie; sie besteht aus den Beiträgen der Z Elektronen des Atoms. Die Bohrschen Postulate sagen aber nichts darüber aus, wie die Elektronen auf die stationären Bahnen zu verteilen sind. Die kleinste Gesamtenergie entsteht, wenn alle Elektronen die kernnächste Bahn besetzen. Dann sollten aber alle Atomeigenschaften monotone Funktionen von Z sein, insbesondere wäre das Periodensystem nicht denkbar. Es muß daher ein *Aufbauprinzip* der Atomhülle geben, das die Besetzung weiter außen liegender Bahnen vorschreibt. Eine Verteilung der Atomelektronen auf die stationären Bahnen nennt man *Elektronenkonfiguration.* Die Periodizität der Atomeigenschaften fordert, daß in der nach wachsenden Kernladungszahlen geordneten Elementereihe jeweils nach Durchlaufen einer Periode Elektronenkonfigurationen auftreten, die hinsichtlich ihres chemischen Verhaltens äquivalent sind.

Das nicht nur für den Atombau grundlegende Aufbauprinzip wurde 1925 von PAULI entdeckt. Es trägt die Bezeichnung *Pauliprinzip* oder *Ausschließungsprinzip.* In seiner für die Atomhülle geltenden ursprünglichen Formulierung fordert es, daß jede stationäre Bahn durch höchstens zwei Elektronen entgegengesetzter Spineinstellung besetzt sein kann.

So ist z. B. die kernnächste Bahn beim Wasserstoffatom mit nur einem, beim Heliumatom mit zwei Elektronen belegt, deren Spins entgegengesetzte Richtungen haben. Beim Lithiumatom besetzt das dritte Elektron, da diese Bahn kein weiteres aufnehmen kann, eine neue, weiter außen liegende stationäre Bahn. Infolgedessen haben Helium und Lithium stark voneinander abweichende chemische Eigenschaften. Ähnlich starke Änderungen der chemischen Eigenschaften treten am Ende einer jeden Periode des Periodensystems auf; nach den chemisch trägen Edelgasen folgen stets die sehr reaktiven Alkalimetalle. Die Elektronenkonfigurationen der Edelgase zeichnen sich durch besondere Stabilität aus, sie bestehen aus *abgeschlossenen Elektronenschalen.* Das zusätzliche Elektron eines Alkaliatoms wird weit außerhalb der abgeschlossenen Schalen des vorhergehenden Edelgases angelagert und verursacht so die Reaktionsfreudigkeit der Alkaliatome. Dieses *Leuchtelektron* erzeugt durch seine Quantensprünge die wasserstoffähnlichen Spektren der Alkaliatome.

Wie die Periodenlängen zeigen, entstehen jeweils nach Anlagerung von 2, 8, 8, 18, 18, 32 Elektronen neue abgeschlossene Schalen. Die Elektronen einer abgeschlossenen Schale besetzen in der Atomhülle stationäre Bahnen, die sowohl energetisch als auch räumlich deutlich über den Bahnen der vorhergehenden Schale angeordnet sind. Die empirisch festgestellten Periodenlängen können im Rahmen der Quantentheorie der Atomhülle begründet werden.

1.1.5. Strukturierter Atomkern

1.1.5.1. Entdeckung der natürlichen Radioaktivität. Nach der Entdeckung der X-Strahlen durch RÖNTGEN vermutete man deren engen Zusammenhang mit Lumineszenzerscheinungen. Es wurden daher die verschiedensten lumineszierenden Mineralien beleuchtet und mit lichtdicht verpackten Fotoplatten in Berührung gebracht. Bei solchen Experimenten entdeckte BEQUEREL 1896, daß Uraniumverbindungen, selbst solche, die keine Lumineszenz zeigen, auch ohne vorherige Beleuchtung die verpackte Fotoplatte schwärzen (Tafel 2). Diese Eigenschaft wurde als *Radioaktivität* bezeichnet.

Marie und Pierre CURIE, die als Assistenten an Bequerels Institut arbeiteten, fanden, daß Pechblende intensiver strahlte als ihrem Uraniumgehalt zukam; sie mußte daher eine wesentlich stärker radioaktive Komponente enthalten. Aus 1 000 kg Pechblenderückständen von der Uraniumgewinnung extrahierten sie 10 kg Bariumsulfat, das etwa 50mal stärker strahlte als Uranium. Der intensive Strahler war daher wie Barium ein Erdalkalimetall und nur schwer von ihm zu trennen. Sie überführten das schwerlösliche Sulfat in das lösliche Chlorid und gewannen durch mehrfaches

Umkristallisieren im Jahre 1898 daraus 0,1 g des Chlorids dieses radioaktiven Metalls, das sie Radium (Ra) nannten. Es strahlte etwa 10^6mal stärker als metallisches Uranium. Reines Radium wurde 1910 von M. CURIE und DEBIERNE elektrolytisch dargestellt. Hinsichtlich der radioaktiven Eigenschaften unterscheidet es sich nicht von seinen Verbindungen. Das von M. und P. CURIE entdeckte Radium emittiert unter allen Bedingungen eine ionisierende Strahlung von etwa 3,3 cm Reichweite in Luft und nahezu gleichbleibender Intensität. Infolge Selbstabsorption dieser Strahlen lumineszieren kompakte Radiumpräparate schwach und geben je Stunde und Gramm Radium mehr als 420 J Wärme an die Umgebung ab.

Neben Radium wurden weitere radioaktive Stoffe entdeckt, so von M. und P. CURIE 1898 das Polonium (Po) als Begleiter von Bismut. MARCKWALD stellte später aus 15000 kg Pechblende 3 mg Polonium her! Fast gleichzeitig mit M. und P. CURIE fand SCHMIDT 1898, daß das seit langem industriell genutzte Element Thorium (Th) ebenfalls radioaktiv ist. 1900 entdeckte DEBIERNE das Actinium (Ac), und DORN stellte durch Ausglühen von Radiumverbindungen fest, daß 1 g Radium immer von 0,6 mm^3 eines radioaktiven Edelgases, das er Radon (Rn) nannte, begleitet ist. RUTHERFORD und SODDY untersuchten das Gas näher. Im Unterschied zum Radium nimmt die Radioaktivität von Radon schon nach wenigen Tagen deutlich ab.

Wenige Jahre nach der Entdeckung durch BEQUEREL waren somit durch die Beiträge von M. und P. CURIE sowie RUTHERFORD und anderen Forschern schon zahlreiche natürlich vorkommende radioaktive Elemente entdeckt. Wegen deren mehr oder weniger schnellen Umwandlung in andere Elemente wurde eine Präzisierung des Elementbegriffs erforderlich. Die Atome eines chemischen Elements mußten nicht mehr unbedingt stabil und unveränderlich sein, sie konnten sich vielmehr in andere Atomsorten umwandeln.

1.1.5.2. Analyse der radioaktiven Strahlung. Die von radioaktiven Körpern emittierte Strahlung besteht im allgemeinen aus verschiedenen Komponenten, die sich hinsichtlich ihrer Beeinflußbarkeit durch elektrische und magnetische Felder sowie ihrer Reichweiten in Materialien unterscheiden. Es treten drei Anteile auf, die nach RUTHERFORD als α-, β- und *γ-Strahlung* bezeichnet werden. α- und β-Strahlen werden im elektromagnetischen Feld wie ein Strom positiv bzw. negativ geladener Teilchen abgelenkt, während die γ-Strahlung davon überhaupt nicht beeinflußt wird. Schon früh erkannte man, daß die Strahlungsarten in der angegebenen Reihenfolge den Kanal-, Katoden- und Röntgenstrahlen äquivalent sind. Aufgrund ihrer Wechselwirkung mit den Atomen wird die Strahlung beim Durchgang durch Gase, Flüssigkeiten oder feste Körper abgeschwächt. Und zwar verhalten sich die durchschnittlichen Reichweiten $\alpha : \beta : \gamma$ etwa wie $1 : 10^2 : 10^4$. Besonders intensiv ionisieren langsame Teilchen mit großer Ladung. Daher sollten α-Teilchen eine größere Ladung und im Mittel eine geringere Geschwindigkeit als β-Teilchen besitzen.

Eine eindrucksvolle Methode, die *Spuren* geladener Teilchen sichtbar zu machen, fand 1911 WILSON mit der *Expansionsnebelkammer*. Ionen können nämlich wie Staubteilchen einem übersättigten Dampf als Kondensationskeime dienen. Wasserdampf wird in der Kammer durch adiabatische Expansion in einen übersättigten Zustand gebracht, und die z. B. von einem α-Teilchen beim Durchlaufen der Kammer erzeugten Ionen bewirken eine Kette von Wassertröpfchen, die den Eindruck einer etwa 0,1 mm breiten Teilchenspur vermitteln.

Nach dem Vorbild der Bestimmung der spezifischen Ladung des Elektrons durch THOMSON führten KAUFMANN und WIECHERT 1901 Ablenkversuche mit β-Strahlen eines Radiumpräparats durch. Sie fanden Teilchengeschwindigkeiten bis zu 97 % der Lichtgeschwindigkeit und eine spezifische Ladung e/m, die für kleine Geschwindigkeiten mit der des Elektrons übereinstimmte, aber für große Geschwindigkeiten weniger als die Hälfte davon betrug. THOMSON zeigte unter Benutzung des Prinzips der Expansionsnebelkammer, daß alle β-Teilchen unabhängig von ihren Geschwindigkeiten die Ladung eines Elektrons tragen und damit tatsächlich Elektronen sind. Die Abnahme der spezifischen Ladung mußte also auf eine entsprechende Zunahme der Teilchenmasse zurückgeführt werden. Vier Jahre vor der Entstehung der relativistischen Mechanik von EINSTEIN existierte somit schon eine ihrer experimentellen Stützen.

RUTHERFORD führte entsprechende Ablenkversuche mit α-Strahlen durch. Er fand Teilchengeschwindigkeiten um 5 % der Lichtgeschwindigkeit. Die gemessene spezifische Ladung ließ den Schluß zu, daß α-Teilchen zwei Elementarladungen und die Masse des Heliumatoms besitzen. α-Teilchen sollten also Heliumatome sein, die zwei Elektronen abgegeben hatten. Diese Hypothese bewiesen RUTHERFORD und ROYDS durch folgendes Experiment: Im Inneren einer total evakuierten Gasentladungsröhre befand sich eine dünnwandige Glasampulle mit Radon. Trotz angelegter Spannung floß zunächst kein Strom, da die Entladungsröhre nur wenige Gasmoleküle enthielt. Nach einiger Zeit jedoch begann der Entladungsraum mit dem für Helium charakteristischen Spektrum zu leuchten. Die vom Radon emittierten α-Teilchen hatten die dünne Glaswand

durchdrungen, dabei zwei Elektronen aufgenommen und als neutrale Heliumatome den Entladungsraum ausgefüllt. Nach Formulierung seines Atommodells im Jahre 1911 konnte Rutherford die Beschreibung der α-Teilchen präzisieren: Es handelt sich um Atomkerne des Heliums.

Die γ-Strahlung wurde von Villard entdeckt und als kurzwellige elektromagnetische Strahlung erkannt. Diese Hypothese fand durch die 1914 von Rutherford und Andrade durchgeführten Beugungsexperimente an Kristallen ihre nachträgliche Bestätigung.

1.1.5.3. Quantitative Beschreibung der Radioaktivität. Rutherford und Soddy erkannten 1903 das Wesen der Radioaktivität in einer spontanen Umwandlung von Atomen, die in jedem Augenblick mit gleicher Wahrscheinlichkeit eintreten kann. Nach der Entdeckung des Atomkerns wurde präzisiert, daß sich dabei die Kerne umwandeln. Als Maß für die *Aktivität* eines Präparats führte man die Einheit 1 Curie (Ci) ein. Ursprünglich sollte 1 Ci die Anzahl der in 1 g Radium je Sekunde ablaufenden Kernumwandlungen sein. Rutherford und Geiger bestimmten erstmals diese Zahl, indem sie die von den α-Teilchen in einem Zinksulfidkristall erzeugten Lichtblitze unter dem Mikroskop auszählten. Der Teilchenstrom wurde durch Blenden auf 1 Teilchen in 2 s reduziert; sie erhielten $1\ \text{Ci} = 3{,}4 \cdot 10^{10}\ \text{s}^{-1}$. 1950 wurde die Aktivität 1 Ci mit $3{,}7 \cdot 10^{10}\ \text{s}^{-1}$ festgesetzt. 1 Ci ist eine sehr große Aktivität. Übliche Präparate besitzen Aktivitäten von einigen Millicurie oder Mikrocurie. Dagegen werden in Cobaltkanonen Aktivitäten der Größenordnung Kilocurie eingesetzt, und die in Kernreaktoren enthaltenen Aktivitäten betragen sogar Megacurie. Im SI-System wird die Einheit Ci durch Bequerel $= 1\ \text{Bq} = 1\ \text{s}^{-1}$ abgelöst.

Untersuchungen, vor allem kurzlebiger radioaktiver Strahler zeigten, daß jede dieser Kernarten eine charakteristische Zeit besitzt, in der sich die Aktivität einer untersuchten Menge der reinen Substanz jeweils halbiert. Diese Zeit wird *Halbwertszeit* genannt. Im Zeitintervall $\mathrm{d}t$ zerfallen, der Rutherfordschen Konzeption entsprechend, $\gamma N\,\mathrm{d}t$ Kerne von insgesamt N Kernen. Dabei ist die *Zerfallskonstante* γ ein Maß für die Zerfallsgeschwindigkeit der Substanz. $\mathrm{d}N = -\gamma N\,\mathrm{d}t$ ist die Änderung (Abnahme) der Zahl der noch nicht zerfallenen Kerne in $\mathrm{d}t$. Diese Differentialgleichung wird durch das Exponentialgesetz $N(t) = N_0 \exp(-\gamma t)$ erfüllt. N_0 ist die Zahl der radioaktiven Ausgangskerne zur Zeit $t = 0$. Mit $N(T) = N_0/2$ findet man für den Zusammenhang zwischen Halbwertszeit T und der Zerfallskonstante: $T = (\ln 2)/\gamma$. Die Zerfallstheorie von Rutherford und Soddy wurde durch die universelle Gültigkeit dieses Zerfallsgesetzes bestätigt.

1.1.5.4. Radioaktive Zerfallsreihen. Ramsay und Soddy füllten ein Gasentladungsrohr mit Radon. Anfangs besaß die Gasentladung das Spektrum von Radon, nach einiger Zeit kamen erwartungsgemäß die Linien von Helium hinzu. Daneben bildete sich aber auf der Innenwand des Entladungsrohrs ein fester „aktiver Niederschlag". Offensichtlich war das Ergebnis des α-Zerfalls von Radon wieder ein radioaktiver Stoff. Also bildet das Uranium der Pechblende mit Radium, Radon und dem weiter zerfallenden „aktiven Niederschlag" eine Kette miteinander verwandter radioaktiver Atomarten, die mit einer stabilen Atomsorte endet. Solche Reihen werden als *radioaktive Zerfallsreihen* oder *radioaktive Familien* bezeichnet.

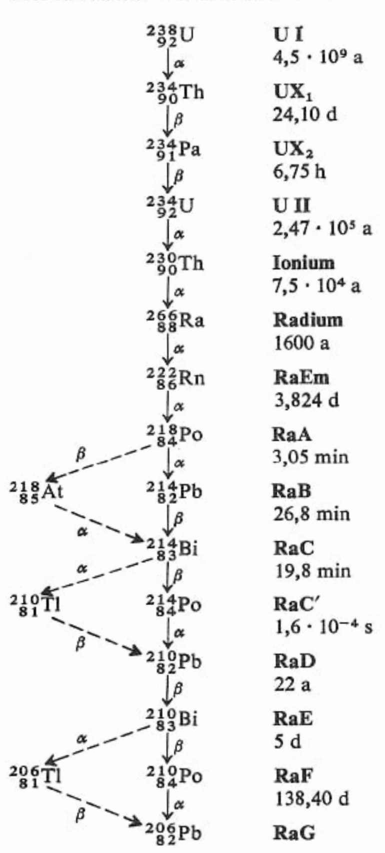


Abb. 1.1.5-1 Uranium-Radium-Reihe; über die gestrichelten Zweige zerfällt der geringere Teil der Kerne; für die Kerne der Hauptlinie sind die historischen Bezeichnungen und die Halbwertszeiten angegeben

Es gibt drei natürlich vorkommende Zerfallsreihen: die *Uranium-Radium-Reihe*, die vom $^{238}_{92}U$ zum $^{206}_{82}Pb$ führt, die *Uranium-Actinium-Reihe* von $^{235}_{92}U$ zu $^{207}_{82}Pb$ und die *Thorium-Reihe* von $^{232}_{90}Th$ zu $^{208}_{82}Pb$. Zur genauen Kennzeichnung der Isotope sind Kernladungszahl und Massenzahl angegeben. Als *Massenzahl* wird die auf eine ganze Zahl gerundete relative Atommasse bezeichnet. Alle früh entdeckten radioaktiven Stoffe gehören zu diesen Reihen. Die Bestimmung der Zerfallsschemata war ein kompliziertes Problem, dessen Lösung nur durch konsequente Anwendung des 1910 von SODDY und MARCKWALD geprägten Begriffs Isotop (↑ 1.1.2.3.) gelang. Die Arbeiten wurden 1913 abgeschlossen, indem SODDY zeigte, daß die Endprodukte der drei Reihen Isotope von Blei mit den Massenzahlen 206, 207 und 208 sind. So ergab sich die Massenzahl des Bleis der Uranium-Radium-Reihe (Abb. 1.1.5-1) unter Berücksichtigung der aus dem Zerfallsschema ersichtlichen acht α-Zerfälle aus der Massenzahl des Uraniums gemäß $238 - 8 \cdot 4 = 206$.

Das erste Glied einer radioaktiven Zerfallsreihe, die *Muttersubstanz*, zerfällt mit großer Halbwertszeit. Alle Zwischenglieder oder radioaktiven *Tochtersubstanzen* sind relativ kurzlebig. Sie entstehen aus der jeweils vorangehenden Substanz und zerfallen in die nachfolgende. Nach hinreichend langer Zeit stellt sich ein *radioaktives Gleichgewicht* ein, d. h., die Bildungsrate der stabilen Endsubstanz ist gleich der Zerfallsrate der Muttersubstanz. Die Teilchenmenge einer radioaktiven Tochtersubstanz ist im Gleichgewicht so groß, daß ihre Zerfallsrate mit der der Muttersubstanz übereinstimmt. Aufgrund des Zerfallsgesetzes erhält man $N_T/T_T = N_M/T_M$, wobei N_T, T_T und N_M, T_M die Teilchenmengen bzw. Halbwertszeiten der Tochter- und Muttersubstanz bedeuten. Nur die Menge des stabilen Endprodukts nimmt ständig zu. Das macht man sich bei der *physikalischen Altersbestimmung* zunutze. So enthalten Gesteine meist Beimengungen von Muttersubstanzen der drei natürlich vorkommenden radioaktiven Zerfallsreihen, die mit Halbwertszeiten von einigen 10^9 a zerfallen. Unter der Voraussetzung, daß seit der Bildung des Gesteins keine Entmischung der Zerfallsprodukte stattgefunden hat, kann sein Alter aus dem Mengenverhältnis von Muttersubstanz und stabilem Endprodukt aufgrund des Zerfallsgesetzes berechnet werden. Auch andere Kernarten werden verwendet, z. B. $^{14}_{6}C$ mit einer Halbwertszeit von etwa $6 \cdot 10^3$ a zur Altersbestimmung abgestorbener biologischer Objekte.

1.1.5.5. Künstliche Kernumwandlung und künstliche Radioaktivität. Natürliche radioaktive Kerne verwandeln sich unter Aussendung von Teilchen in andere Kerne. Der umgekehrte Vorgang, die Herstellung eines Kerns aus einem anderen durch Teilchenbeschuß, ist eine *künstliche Kernumwandlung*. Falls der entstehende Kern instabil ist und daher radioaktiv zerfällt, spricht man von *künstlicher Radioaktivität*. Zur Zeit der Entdeckung dieser Prozesse standen für den Kernbeschuß vor allem die von radioaktiven Stoffen emittierten α-Teilchen und Elektronen zur Verfügung. Elektronen erwiesen sich als unwirksam, sie wurden vom Kern bei den damals verfügbaren Energien nur elastisch reflektiert, und α-Teilchen müssen schon ziemlich hohe Geschwindigkeiten haben, um trotz der elektrostatischen Abstoßung durch den Kern diesem genügend nahe zu kommen.

Die erste künstliche Kernreaktion gelang 1919 RUTHERFORD unter Verwendung der energiereichen α-Strahlung von $^{214}_{84}Po$, damals als Ra C′ bezeichnet. Er ließ die α-Strahlen durch Stickstoffgas treten und beobachtete in einem dahinter angebrachten Zinksulfidkristall neben den starken Lichtblitzen der α-Teilchen schwächere, die eindeutig auf Protonen zurückgeführt werden konnten. Diese mußten aufgrund der Reaktion $^{14}_{7}N + ^{4}_{2}He \rightarrow ^{17}_{8}O + ^{1}_{1}H$ entstanden sein. 1925 wiederholte BRACKETT das Experiment in einer Wilsonschen Nebelkammer und fand die Reaktion als „Stern" dreier Teilchenspuren unterschiedlicher Dicken. In den folgenden Jahren wurden entsprechende Umwandlungen von Bor, Silicium und Kalium in stabile Isotope von Kohlenstoff, Phosphor und Calcium durchgeführt.

Ein radioaktiver Kern konnte erstmals 1934 von I. JOLIOT-CURIE und F. JOLIOT erzeugt werden. Sie benutzten die α-Strahlung eines Poloniumisotops und erhielten durch Beschuß von $^{27}_{13}Al$ den Kern $^{30}_{15}P$, der mit der Halbwertszeit 2,5 min in $^{30}_{14}Si$ zerfällt. Neben den mehr als 40 natürlich vorkommenden radioaktiven Kernarten war dies der erste einer heute auf weit über 1000 angewachsenen Reihe künstlich hergestellter radioaktiver Kerne.

Künstliche *radioaktive Isotope* der verschiedensten Elemente haben in Wissenschaft und Technik eine große Zahl von Einsatzbereichen erschlossen.

1.1.5.6. Teilchenbeschleuniger. Die Verwendung von Teilchen aus natürlichen radioaktiven Quellen für Experimentierzwecke ist mit Nachteilen verbunden. Die Maximalenergie ist vorgegeben und die Strahlintensitäten sind klein, denn die Strahlung wird isotrop emittiert und vom kompakten Präparat größtenteils absorbiert. Beispielsweise sind von den etwa $3{,}6 \cdot 10^{10}$ α-Teilchen, die in 1 g Radium je Sekunde entstehen, im Strahl nur etwa 10^4 nutzbar. Es bestand daher ein dringendes Bedürfnis nach steuerbaren Quellen hoher Intensität. *Teilchenbeschleuniger* sind solche Quellen. Diese Anlagen beschleunigen geladene

Teilchen im Vakuum mit Hilfe elektrischer Felder. In einem ersten Schritt muß daher stets ein genügend intensiver Strom von Ionen hergestellt werden. Anfangs benutzte man Kanalstrahlrohre, gefüllt z. B. mit Wasserstoff oder Helium, zur Erzeugung eines Stroms von Protonen bzw. α-Teilchen. Stromstärken von 100μA, das sind einige 10^{14} Teilchen je Sekunde, ließen sich so erreichen. Der zweite Schritt ist der eigentliche Beschleunigungsvorgang.

1932 führten COCKCROFT und WALTON erstmals eine Kernreaktion unter Verwendung künstlich beschleunigter Teilchen herbei. Ein im Kanalstrahlrohr erzeugter und beschleunigter Protonenstrom wurde auf schwere Kerne geleitet und verursachte deren „*Zersplitterung*" in mehrere leichtere Kerne. Zur Beschleunigung auf höhere Energien baute VAN DE GRAAFF 1933 einen *Bandgenerator*, der etwa 10^6 V erzeugte. Ladung wurde auf ein umlaufendes Isolatorband gesprüht, das diese ins Innere einer hohlen Metallkugel mit ungefähr 1 m Durchmesser transportierte. Die Ladung sammelte sich auf der Oberfläche dieser Hochspannungskugel. Einfach geladene Teilchen konnten damit auf eine Energie von 1 MeV gebracht werden. Die Durchschlagfestigkeit des Gases, in das die Hochspannungskugel eingebettet ist, begrenzt auch heute die Leistungsfähigkeit solcher Anlagen auf einige 10^6 V.

Höhere Energien erreicht man, indem das erzeugte elektrische Feld mehrfach von den Teilchen durchlaufen wird. Je nach der Form der Teilchenbahn unterscheidet man *Linearbeschleuniger* und *Kreisbeschleuniger*.

Nach Vorarbeiten von WIDEROE (1928) bauten SLOAN und LAWRENCE 1931 einen Hochfrequenz (HF)- Linearbeschleuniger. Das 30 m lange Gerät bestand in einer linearen Anordnung von 30 Metallröhren mit in Ausbreitungsrichtung der Ionen wachsenden Längen. Die Pole der Hochspannungsquelle (45 kV, 7 MHz) wurden in wechselnder Folge mit den Röhren verbunden. Eine Ionenwolke bewegt sich dann innerhalb der Röhren kräftefrei und trifft zwischen den Röhren dank deren geeignet gewählten Längen stets auf ein beschleunigendes Feld. Hg-Ionen konnten so auf eine Energie von 1,26 MeV gebracht werden. Moderne Linearbeschleuniger arbeiten nach dem Wanderwellenprinzip und liefern Elektronen von einigen 10 GeV.

Den Ausgangspunkt für die Entwicklung der *Kreisbeschleuniger* bildete das 1932 von LAWRENCE konstruierte *Zyklotron* (Tafel 3). Zwei HF-Elektroden, die D-förmig sind und als Duanden bezeichnet werden, sind von einem zeitlich konstanten, homogenen Magnetfeld senkrecht durchsetzt. In jeder Schwingungsdauer der HF-Spannung wird eine Ionenwolke in der Mitte emittiert, im Spalt zwischen den Elektroden beschleunigt und innerhalb der Elektroden unter dem Einfluß des Magnetfelds jeweils auf einem Halbkreis zum Spalt zurückgeführt. Nach etwa 100 Umläufen besitzen die Ionen Bewegungsenergien um 1% ihrer Ruhenergie; für Protonen sind das 10 MeV. Höhere Energien lassen sich nicht erreichen, denn wegen der Geschwindigkeitsabhängigkeit der Ionenmasse geraten die umlaufenden Ionen gegenüber der konstanten Frequenz der HF-Spannung außer Takt, wenn die Ionengeschwindigkeit in die Nähe von 10% der Lichtgeschwindigkeit kommt. Als Elektronenbeschleuniger eignet sich das Zyklotron daher überhaupt nicht; die Elektronen würden nur eine Energie von etwa 5 keV erhalten. Durch geeignete Variation der Frequenz der Beschleunigungsspannung, durch Verwendung zeit- und ortsabhängiger Magnetfelder sowie durch Einführung supraleitender Magnetspulen erreichen moderne Kreisbeschleuniger Protonenenergien von einigen Hundert GeV.

Als weniger kostspielige Anlage zur Beschleunigung von Elektronen, vor allem für technische Zwecke, hat sich das 1941 von KERST nach einer Idee von WIDEROE (1927) konstruierte *Betatron* bewährt. Es arbeitet nach dem Transformatorprinzip, d. h., von einem zeitabhängigen magnetischen Feld wird ein elektrisches Wirbelfeld erzeugt, in dem die Elektronen nach etwa 10^5 Umläufen im allgemeinen Energien bis zu 30 MeV erreichen. Bei höheren Energien macht die Strahlungsdämpfung der Elektronen den Betrieb des Gerätes unwirtschaftlich.

Teilchenbeschleuniger erweiterten die Möglichkeiten der experimentellen Kernphysik wesentlich (Tafeln 4/5). So wurden in den Jahren 1937, 1939 und 1940 mit der Herstellung der Kerne von Technetium ($^{97}_{43}$Tc), Francium ($^{223}_{87}$Fr) bzw. Promethium ($^{145}_{61}$Pm) und Astat ($^{210}_{85}$At) im Periodensystem verbliebene Lücken geschlossen. Die Synthese des ersten Transurans Neptunium ($^{237}_{93}$Np) leitete eine Entwicklung ein, die bis 1980 bis zum Element mit der Kernladungszahl 107 geführt hat.

1.1.5.7. Entdeckung des Neutrons. RUTHERFORD vermutete schon 1920, daß es einen Kern mit der Massenzahl 2 und der Kernladungszahl 1 geben müsse. Dieses als *Deuterium* bezeichnete Wasserstoffisotop wurde 1932 von UREY und seinen Mitarbeitern gefunden. Sein Kern, das Deuteron, muß neben einem Proton ein Teilchen ohne Ladung mit der Massenzahl 1 enthalten.

BOTHE und BECKER experimentierten 1930 mit α-Strahlung von 5,3 MeV und leichten Kernen wie Li, Be und B. Sie stellten eine „harte, durchdringende" Strahlung fest, die durch 5 cm Blei nur zur Hälfte absorbiert wurde. Darin verhielt sie sich wie γ-Strahlung von 5 MeV. I. JOLIOT-CURIE und F. JOLIOT zeigten

1931, daß diese Strahlung aus wasserstoffhaltigen Substanzen sogenannte „Rückstoßprotonen" von 4,5 MeV Energie ablöst. Solche Prozesse konnte γ-Strahlung nur bei einer Energie von 50 MeV hervorrufen. Die Interpretation dieser Komponente als γ-Strahlung wurde völlig unhaltbar, als es CHADWICK 1932 gelang, „Rückstoßkerne" in schweren Gasen bis Krypton ($^{86}_{36}$Kr) zu erzeugen. Unter Benutzung von Energie- und Impulserhaltungssatz wies er nach, daß es sich bei der untersuchten Strahlung nur um einen Strom von Teilchen mit etwas mehr als Protonenmasse handeln konnte.

Das neue Teilchen besitzt keine Ladung und wurde *Neutron* ($^{1}_{0}$n) genannt. Es entsteht z. B. bei α-Bestrahlung von Beryllium aufgrund der Reaktion $^{9}_{4}\text{Be} + ^{4}_{2}\text{He} \rightarrow ^{12}_{6}\text{C} + ^{1}_{0}\text{n}$. Wie Proton und Elektron besitzt es den Spin 1/2. Obwohl ohne Ladung, hat das Neutron ein entgegen dem Spin gerichtetes magnetisches Eigenmoment. Proton und Neutron werden als Nukleonen bezeichnet und sind die Bausteine der Atomkerne. Durch diese Schlußfolgerung begründeten 1932 HEISENBERG und IWANENKO die Kerntheorie. Ein Kern der Massenzahl A und der Kernladungszahl Z enthält Z Protonen und A–Z Neutronen. Außerhalb des Kerns ist das Neutron instabil und zerfällt mit einer Halbwertszeit von 10,8 min in ein Proton, ein Elektron und ein Antineutrino. Aufgrund seiner Ladungslosigkeit tritt das Neutron nicht in elektrostatische Wechselwirkung mit den Bestandteilen der Atome. Daraus resultiert einerseits das hohe Durchdringungsvermögen der Neutronen, und andererseits sind schon sehr langsame Neutronen fähig, Kernreaktionen einzuleiten.

1.1.5.8. Kernspaltung und ihre Anwendung. Bei der *Kernspaltung* zerfällt ein Kern spontan oder induziert in zwei, seltener auch drei nahezu gleich große Teile. Anfangs hielt man eine solche Reaktion für unmöglich. Daher war die Vorgeschichte der Entdeckung der Kernspaltung durch eine andere Zielstellung gekennzeichnet.

1934 zeigte FERMI durch Neutronenbeschuß verschiedener Materialien, daß ein Kern unter Aufnahme eines Neutrons in einen isotopen Kern oder, falls ein zusätzlicher β-Zerfall stattfindet, in einen Kern des im Periodensystem folgenden Elements übergeht. Insbesondere die Neutronenreaktion mit Uranium lieferte zahlreiche radioaktive Reaktionsprodukte. FERMI nahm an, daß bei dieser Reaktion zunächst Transurane entstanden seien. Radioaktive Zerfallsreihen konnten dann eine große Zahl strahlender Endprodukte mit großen Kernladungszahlen erzeugen. Auch HAHN, MEITNER und STRASSMANN waren zunächst dieser Ansicht, nachdem sie die Experimente 1936/37 sorgfältig wiederholt hatten (Tafel 2). Die Wende trat ein, als I. JOLIOT-CURIE und SAVITCH unter den Zerfallsprodukten das Element Lanthan fanden. HAHN und STRASSMANN stellten 1938 fest, daß auch Barium zu den Reaktionsprodukten gehört. Die Vermutung einer durch Neutronenbeschuß induzierten *Kernspaltung* des Uraniums bestätigte sich, als es ihnen gelang, auch das jeweils andere Spaltprodukt, Brom bzw. Krypton, nachzuweisen. Neben diesen beiden zuerst gefundenen Zerfallsvarianten gibt es andere; die Spaltung in zwei gleich große Teile ist aber wenig wahrscheinlich. MEITNER und FRISCH erklärten die Kernspaltung anhand des Modells eines schwingenden geladenen Flüssigkeitstropfens und machten 1939 erstmals auf den energetischen Aspekt der Spaltung schwerer Kerne aufmerksam. Da sich die Spaltprodukte unmittelbar nach ihrer Trennung in Abständen der Größenordnung 10^{-14} m befinden, müssen sie aufgrund ihrer elektrostatischen Abstoßung hohe Geschwindigkeiten erreichen. Sie berechneten, daß je Spaltungsakt von $^{235}_{92}$U etwa 170 MeV Bewegungsenergie der Bruchstücke erzeugt werden. Dagegen läuft die Verbrennung von Kohle als chemische Reaktion in der Atomhülle mit einem Energieumsatz von einigen Elektronenvolt je Oxydationsakt ab. Dementsprechend liefert die Spaltung von 1 kg U-235 etwa die $2{,}5 \cdot 10^6$fache Energiemenge wie die Verbrennung von 1 kg Steinkohle.

Verglichen mit makroskopischen Energiemengen ist die je Spaltungsakt freigesetzte Energie dennoch verschwindend klein. Die Kernenergie wird erst nutzbar, wenn man je Zeiteinheit eine genügend große Anzahl von Spaltungen herbeiführen kann. Das wiederum setzt Neutronenmengen voraus, die sich nicht von außen bereitstellen lassen. 1939 machten JOLIOT, KOWARSKI und HALBAN darauf aufmerksam, daß durch die Spaltung selbst je Spaltungsakt im Mittel 2 bis 3 Neutronen entstehen. Wenn also genügend viele Neutronen wieder auf spaltbare Kerne treffen, so kann eine *Kettenreaktion* entstehen. Ausgehend von dieser Vorstellung baute FERMI in Chicago den ersten *Kernreaktor* auf, einen „Haufen" von Uranium und Uraniumoxid, der am 2. 12. 1942 kritisch wurde, d. h. durch Temperaturerhöhung anzeigte, daß eine kontrollierte Kettenreaktion angelaufen war. Unter bestimmten Bedingungen nimmt die Kettenreaktion einen unkontrollierten Verlauf. Darin besteht das Wirkprinzip der Atombombe.

Schon HAHN und STRASSMANN fanden, daß neben Uranium auch Thorium durch Neutronenbeschuß zur Kernspaltung angeregt werden kann. Heute kennt man eine Vielzahl von Kernen, die sich ähnlich verhalten, wobei manche wie U-235 bereits empfindlich auf langsame Neutronen reagieren und andere wie U-238 nur auf schnelle.

Die Kernspaltung läuft auch spontan ohne äußere Einwirkung ab. Jedoch betragen die Halbwertszeiten für U-238 und Th-232 einige 10^{16} a, erst die höheren Transurane neigen wesentlich stärker zur spontanen Spaltung. Im Jahre 1940 zeigte es sich, daß entsprechend den Vorstellungen von FERMI unter den Reaktionsprodukten auch *Transurane* enthalten sind. Neptunium und später auch andere wurden nachgewiesen. Heute produzieren moderne Kernreaktoren auf der Basis von U-235 große Mengen von Pu-239, das durch Neutroneneinfang aus U-238 und einem nachträglichen β-Zerfall entsteht. Dieses Plutoniumisotop ist unter ähnlichen Bedingungen wie U-235 spaltbar und daher selbst ein Kernbrennstoff (Tafeln 6/7).

1.1.5.9. Kernfusion. Die Erkenntnisse über die Kernstruktur und die Möglichkeit von Kernumwandlungen führen zu dem Schluß, daß neben der Spaltung schwerer auch die Verschmelzung leichter Kerne auftreten kann, wobei ebenfalls Energie frei gesetzt wird. Diese Prozesse werden unter irdischen Bedingungen normalerweise durch die Abstoßung zwischen den elektrisch geladenen Kernen verhindert. Sie laufen aber ständig im Inneren von Sternen bei Temperaturen der Größenordnung 10^7 K ab, insbesondere im Sonneninneren. Dabei verschmelzen über verschiedene Zwischenstufen je 4 Protonen zu einem Heliumkern. Diese Fusionsprozesse sind die Quelle der Sonnenenergie und damit eine Voraussetzung des Lebens auf der Erde und Primärquelle fast all unserer Energieressourcen. Bei noch höheren Temperaturen können etwa durch Verschmelzung dreier Heliumkerne Kohlenstoffkerne gebildet werden. Die kernphysikalischen Erkenntnisse gestatten zusammen mit astrophysikalischen Daten und theoretischen Modellen über die Vorgeschichte des Universums, den Prozeß der Elemententstehung im Detail zu analysieren und die beobachtete Häufigkeitsverteilung der Elemente zu erklären.

In der *Wasserstoffbombe* wird eine unkontrollierte Fusionsreaktion ausgelöst, wobei die dazu erforderliche hohe Temperatur durch eine Kernspaltungsbombe erzeugt wird (Tafel 17). Die vom Menschen kontrollierte und gesteuerte Kernfusion wird die fast unerschöpfliche Energiequelle der Zukunft sein. Dabei kommt als „Brennstoff" Deuterium, also schwerer Wasserstoff H–2, in Betracht. Das im Wasser der Ozeane vorhandene schwere Wasser könnte den Energiebedarf der Menschheit für Hunderte von Millionen Jahren decken. Gegenwärtig werden zwei Wege zur Realisierung der gesteuerten Fusion beschritten: die thermonukleare Reaktion in heißen Plasmen und die laserinduzierte Fusion, bei der kleine Tröpfchen von Fusionsmaterial durch extrem intensive Laserstrahlung aufgeheizt werden.

1.1.6. Weg zur Elementarteilchenphysik

1.1.6.1. Situation vor 1950. Die Welt der bekannten Elementarteilchen war besonders einfach im Jahre 1932, als man nach der Entdeckung des Neutrons (↑ 1.1.5.7.) den Aufbau der Atomkerne verstand. Außer dem Lichtquant, das eine gewisse Sonderstellung auch in der modernen Elementarteilchenphysik behalten hat, bestand die materielle Welt anscheinend aus nur 3 Teilchen: dem Elektron und den beiden etwa 2000mal schwereren Nukleonen, dem Proton und dem Neutron. Nur langsam und unzusammenhängend kamen vereinzelte Hinweise auf weitere Teilchen. 1930 entwickelte DIRAC eine relativistische Quantenmechanik des Elektrons, die ihn zu der Voraussage des *Positrons* führte, des „Zwillingsbruders" des Elektrons, das sich von diesem bei gleicher Masse und gleichem Spin nur durch das positive Ladungsvorzeichen unterscheiden sollte. Bereits 1932 wurde dieses Antiteilchen des Elektrons mit den vorausgesagten Eigenschaften als Bestandteil der kosmischen Strahlung nachgewiesen. Die weitere Entwicklung der Quantenfeldtheorie führte zu dem Schluß, das zu jedem Teilchen auch ein Antiteilchen gehören müsse, insbesondere sollte es Antiprotonen und Antineutronen geben. Die Suche nach diesen Teilchen in der Höhenstrahlung verlief jedoch ergebnislos, und die verfügbaren Beschleuniger lieferten zunächst noch keine Teilchen mit Energien im GeV-Bereich, die zur Erzeugung von Nukleonen erforderlich sind. Erst 1955 gelang in Berkeley (Kalifornien) die Erzeugung von Proton-Antiproton-Paaren beim Beschuß von Materie mit energiereichen Protonen, 1956 wurden am gleichen Beschleuniger Antineutronen durch eine Art Ladungsneutralisation von Proton-Antiproton-Paaren erzeugt. Ebenso wie Elektron-Positron-Paare kann jedes Teilchen zusammen mit seinem Antiteilchen zerstrahlen (s. a. 1.2.2.4.) und dabei Photonen oder beliebige andere Teilchen-Antiteilchen-Paare bilden. Durch die Existenz von Antiteilchen kommt eine wesentliche Symmetrie in die Welt der Elementarteilchen, ohne daß damit eine prinzipielle Erweiterung der Mannigfaltigkeit von Teilchen verbunden wäre.

Noch zwei weitere Hinweise auf völlig neue Teilchen ergaben sich aus der Kernphysik der 30er Jahre: 1931 postulierte PAULI, daß die scheinbare Verletzung des Energiesatzes beim β-Zerfall von Atomkernen auf die Emission eines sehr leichten ungeladenen Teilchens, des *Neutrinos*, zurückzuführen sei. Dieses Teilchen steht mit anderen nur in einer so schwachen Wechselwirkung, daß es selbst den Erdball mit hoher Wahrscheinlichkeit ungestört passieren kann. Daher wurde seine Existenz lange nur indirekt durch die Analyse des β-Zerfalls, ein-

schließlich der Messung des Rückstoßes des Endkerns, belegt. Erst 1956 gelang es COWAN und REINES mit außerordentlich großem apparativem Aufwand, Neutrinos (genauer Elektron-Antineutrinos) direkt durch eine zum β-Zerfall inverse Reaktion nachzuweisen.
Die Natur der die Nukleonen zusammenhaltenden Kernkräfte blieb zunächst im Dunklen. Nach den allgemeinen Erkenntnissen der Quantenfeldtheorie sollten einem Kraftfeld wiederum Teilchen als Feldquanten zugeordnet sein. Die Wechselwirkung wird dann durch den Austausch solcher Quanten zwischen den ursprünglichen Teilchen vermittelt. Die elektromagnetischen Kräfte entstehen in diesem Bild durch den Austausch von Photonen zwischen geladenen Teilchen. 1935 postulierte YUKAWA die Existenz der Quanten des Kernkraftfelds. Er nannte sie *Mesonen* [*mesos*, griech. das mittlere], da ihre Masse mit einigen Hundert Elektronenmassen zwischen Elektronen- und Nukleonenmasse liegen sollte, um die Reichweite der Kernkräfte von etwa 10^{-15} m zu erklären. Tatsächlich fand man 1938 in der kosmischen Strahlung ein Teilchen mit der Masse $207m_{el}$, das aber wegen seiner geringen Wechselwirkung mit Kernen nicht das von YUKAWA vorausgesagte sein konnte. Es ist das heute als *Myon* bezeichnete Teilchen, das sich vom Elektron (bzw. bei positiver Ladung vom Positron) nur durch seine größere Masse unterscheidet. Das als *π-Meson* oder *Pion* bezeichnete Yukawaquant wurde erst 1947 ebenfalls in der kosmischen Strahlung entdeckt. Sowohl Myon wie Pion sind auch im Vakuum instabil, sie zerfallen mit sehr unterschiedlichen mittleren Lebensdauern in leichtere Teilchen. Die Existenz des *Myons*, das etwa 10^{-6} s lebt, erschien (und erscheint noch) als eine völlig überflüssige Laune der Natur, die übrigen Teilchen hatten definierte Funktionen beim Aufbau der uns umgebenden Materie und bei Kernumwandlungen (Neutrino).

1.1.6.2. Höhere Energien – mehr Elementarteilchen. Mit der Entwicklung immer leistungsfähigerer Teilchenbeschleuniger wuchs in den 50er Jahren die Zahl der entdeckten Teilchen rasch an; so fand man die *Kaonen* mit etwa halber und die *Hyperonen* mit ein- bis zweifacher Nukleonenmasse. Die Hyperonen zerfallen in etwa 10^{-10} s. Es entstand eine fast verwirrende Vielzahl von Reaktionen und Zerfällen. 1960 waren insgesamt etwa 30 Elementarteilchen bekannt, das einfache Bild des Jahres 1932 war völlig zerstört. Mit der weiteren Entwicklung der Beschleunigertechnik kamen ständig neue, schwerere Teilchen hinzu. Die 60er Jahre brachten die sogenannten *Resonanzen*, das sind Teilchen mit einer so extrem kurzen Lebensdauer von etwa 10^{-23} s, daß sie sich jedem direkten Nachweis, etwa als Spur in einer Blasenkammer, entziehen und ihre vorübergehende Existenz nur in einer resonanzartigen Energieabhängigkeit der Wirkungsquerschnitte für Reaktionen zwischen langlebigen Teilchen verraten, nämlich dann, wenn die Resonanzen als Zwischenprodukte gebildet werden.
Weit entfernt von einem tieferen theoretischen Verständnis wurden Erfolge bei der Klassifizierung der Teilchen und der zwischen ihnen wirksamen Wechselwirkungen erreicht. Völlig neue ladungsartige Kenngrößen der Teilchen wurden eingeführt, um die bei der systematischen Durchmusterung des rasch anwachsenden experimentellen Materials deutlich werdenden Reaktionsverbote durch Erhaltungssätze für diese neuen „Ladungen" zu erfassen. Sehr erfolgreich war ein 1961 von GELLMANN und NÉEMANN eingeführtes, als unitäre Symmetrie bezeichnetes Klassifizierungsschema, dessen heuristische Bedeutung mit dem Periodensystem der Elemente verglichen werden kann. Auf theoretische Einsicht in die Wechselbeziehungen zwischen den Elementarteilchen gegründet, gelang die Voraussage eines Teilchens mit all seinen wesentlichen Eigenschaften. Dieses Teilchen, das *Ω^--Hyperon*, wurde 1964 tatsächlich gefunden (Tafel 9).
Das neue Klassifizierungsschema legte es nahe, die *Hadronen*, das sind alle Teilchen außer Lichtquant, Elektron und Positron, positiven und negativen Myonen und den Neutrinos, aus nur 3 fundamentaleren Teilchen und ihren Antiteilchen aufgebaut zu denken. Diese Teilchen nannte man Quarks, sie müßten elektrische Ladungen $\pm e/3$ und $\pm 2e/3$ haben. Diese Teilchen konnten bisher nicht als freie Teilchen nachgewiesen werden, möglicherweise können sie als solche überhaupt nicht existieren. Neuere vielversprechende theoretische Ansätze beschreiben die Hadronen als gebundene Zustände von Quarks, die durch den Austausch von sogenannten *Gluonen* untereinander wechselwirken. Falls sich dieses Konzept als richtig erweist, so würde die Theorie der Hadronen auf einem neuen fundamentaleren Niveau wieder ein bemerkenswert einfaches Grundrezept realisieren.
1974 wurden bei Experimenten an neuen leistungsfähigen Elektron-Positron-Speicherringen mit gegenläufigen Strahlen neue Teilchen gefunden, die trotz ihrer großen Masse eine ungewöhnlich lange Lebensdauer haben. Auch diese überraschende Entdeckung zeigt, wie sehr die Elementarteilchenphysik ein Gebiet ohne absehbaren Abschluß ist.
Das Hinabsteigen zu immer tieferen materiellen Strukturen hat auch einen wesentlichen kosmologischen Aspekt. Die frühesten Phasen der Entwicklung des Universums sind durch so hohe Temperaturen charakterisiert, daß vor der Bildung der Atomkerne nur Elementarteilchen

existieren, und zwar auch bzw. sogar ausschließlich solche, die heute nur in gigantischen Beschleunigern erzeugt oder überhaupt noch nicht erzeugt oder nur als seltene Ereignisse in der kosmischen Strahlung registriert werden können. Fortschritte in der Analyse der frühesten Jugend des Universums sind an die Erkenntnisse der Elementarteilchenphysik gebunden.

1.2. Welle-Teilchen-Dualismus und Plancksches Wirkungsquantum

1.2.1. Teilchen und Wellen in der klassischen Physik

Die Begriffe *Teilchen* und *Welle* sind grundlegend für die klassische Physik. So kann z. B. ein makroskopischer Körper für die Formulierung der Bewegungstheorie durch ein Teilchen ersetzt werden, sofern nur die Bewegung seines Massemittelpunkts interessiert. Dagegen bilden Ausbreitungsvorgänge ähnlich den Oberflächenwellen von Flüssigkeiten die Grundlage des Wellenbegriffs der klassischen Physik. Mit diesen Begriffen verbinden sich daher im Rahmen der klassischen Physik entstandene Vorstellungen von einem physikalischen Objekt bzw. einer speziellen Bewegung.
Die *typischen Eigenschaften von Teilchen* sind *Individualität* und *Lokalisierbarkeit*, d. h., ein Teilchen tritt immer als ganzes Teilchen in Erscheinung, und sein Ort kann zu jedem Zeitpunkt mit genügender Genauigkeit gemessen werden. Die Messung des Ortes, z. B. mit Hilfe von Licht, ist wegen der großen Teilchenmasse praktisch ohne Rückwirkung auf die Teilchenbewegung. Alle gemessenen Teilchenorte bilden eine glatte Raumkurve, die Bahn des Teilchens.
Physikalische Größen, deren Schwingungszustände sich in Form von *Wellen* ausbreiten, besitzen *kontinuierliche Wertevorräte*. Die störungsfreie Überlagerung von Wellen wird als *Interferenz* bezeichnet.
In der klassischen Physik ist eine physikalische Erscheinung in jedem Fall Teilchen *oder* Welle, je nachdem, ob aus Experimenten Individualität und Lokalisierbarkeit oder Kontinuität und Interferenz folgen. Besonders deutlich werden die Unterschiede beim *Doppelspaltexperiment* sichtbar, das erstmals 1803 von Young als Modell für Beugungsexperimente mit Licht beschrieben wurde. Die zu untersuchende Strahlung fällt auf eine undurchlässige Wand mit zwei parallelen Spalten (Abb. 1.2.1-1). Gemessen wird die Intensität der auf dem Schirm einfallenden Strahlung.
Führt man das *Experiment mit Teilchen* durch (↑ Abb. 1.2.1-1a), so treffen auf dem Schirm stets ganze Teilchen ein, niemals Bruchstücke. Jedes auf dem Schirm nachgewiesene Teilchen muß durch einen bestimmten Spalt getreten sein, denn Teilchen bewegen sich auf wohldefinierten Bahnen. Bei zwei geöffneten Spalten beobachtet man daher die Summe der Verteilungen, die bei wechselseitigem Schließen jeweils eines Spalts entstehen.
Fällt eine Welle auf den Doppelspalt (↑ Abb. 1.2.1-1 b), so erzeugen die von beiden Spalten ausgehenden Elementarwellen ψ_1 und ψ_2 auf dem Schirm durch Interferenz die typische Beugungsfigur mit der kontinuierlichen Intensitätsverteilung $|\psi_1 + \psi_2|^2 = |\psi_1|^2 + 2\,\mathrm{Re}(\psi_1\psi_2^*) + |\psi_2|^2$. Diese unterscheidet sich durch den Interferenzterm $2\,\mathrm{Re}(\psi_1\psi_2^*)$ von der Summe der für jeweils einen offenen Spalt gefundenen Verteilungen, und das Doppelspaltexperiment mit Wellen führt zu einem grundsätzlich anderen Ergebnis als mit Teilchen.

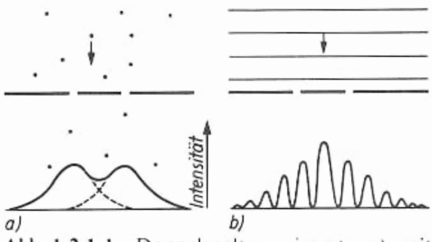


Abb. 1.2.1-1 Doppelspaltexperiment; a) mit Teilchen, b) mit Wellen

1.2.2. Lichtquanten

1.2.2.1. Zur Entwicklung der Theorie des Lichts. Die experimentelle Erkundung der Natur des Lichts begann im 17. Jahrhundert. Sie wurde mit dem im Rahmen der klassischen Physik einzig möglichen Ziel geführt, Licht entweder als Teilchen- oder als Wellenphänomen zu verstehen. Snellius und Descartes beschrieben 1637 Reflexion und Brechung von Licht (↑ 2.3.1.7.), Grimaldi entdeckte 1665 die Beugung von Licht, und Römer bestimmte 1676 aus der Verfinsterung der Jupitermonde die Zeit, in der das Licht den Durchmesser der Erdbahn durchläuft, und damit die Vakuumlichtgeschwindigkeit.
Erste theoretische Vorstellungen von der Natur des Lichts wurden 1675 von Newton und 1690 von Huygens veröffentlicht. Die *Emissionstheorie* von Newton beschrieb Licht als einen Strom sehr kleiner, von der Lichtquelle emittierter Teilchen. Unterschiedliche Teilchengrößen sollten die verschiedenen Farben des Lichts hervorbringen. Huygens begründete die *Wellentheorie des Lichts*. Er betrachtete Licht als wellenförmige Ausbreitungserscheinung gewisser Ätherschwingungen, die durch

entsprechende Schwingungen der Teilchen leuchtender Körper angeregt wurden. Obwohl die Teilchentheorie für die teilweise Reflexion von Licht nur eine unbefriedigende und für die Beugung gar keine Erklärung geben konnte, herrschte im 18. Jahrhundert wegen der Autorität von NEWTON die Teilchenauffassung vor.

Zu Beginn des 19. Jahrhunderts wurden Interferenz und Beugung (↑ 2.3.2.) ausführlich von YOUNG untersucht und aufgrund der Wellentheorie erklärt. MALUS entdeckte 1809, daß Licht durch Reflexion in einen polarisierten Zustand überführt werden kann, in dem nicht mehr alle Richtungen senkrecht zur Ausbreitungsrichtung des Lichts gleichberechtigt sind. So konnte FRESNEL durch seine *Äthertheorie des Lichts* die Huygenssche Wellentheorie präzisieren. Licht ist danach eine transversale Schwingung des elastischen Äthers, die sich wellenförmig ausbreitet. FRAUNHOFER und FRESNEL brachten die Theorie der Beugung von Licht in eine mathematische Form. Für die Lichtgeschwindigkeit wurden durch Experimente von FIZEAU (1849, 1851), FOUCAULT (1850) und später von MICHELSON (ab 1878) genauere Werte bestimmt. Ihre Vollendung fand die Wellentheorie des Lichts 1861 in der *elektromagnetischen Lichttheorie* von MAXWELL. Diese beschreibt Licht als eine transversale elektromagnetische Welle, die man sich zunächst als Schwingung eines Lichtäthers vorstellte. 1881 führte MICHELSON erstmals sein berühmtes Experiment durch, dessen Ziel der Nachweis der Bewegung der Erde relativ zum Äther war. Eine solche Relativbewegung konnte nicht nachgewiesen werden, auch nicht von JOOS 1930 bei höchstpräziser Durchführung des Experiments. Es wurden mehrere Versuche unternommen, dieses Ergebnis unter Aufrechterhaltung der Ätherhypothese zu erklären. Aber nur die 1905 von EINSTEIN gegebene Interpretation, daß es keinen Äther gibt und daß die Lichtausbreitung in allen Bezugssystemen isotrop mit gleicher Geschwindigkeit erfolgt, erwies sich als tragfähig. Im 19. Jahrhundert dominierte die Wellentheorie.

Den Ausgangspunkt für die Vereinigung von Teilchen- und Wellenaspekt bei der Beschreibung der Natur des Lichts bildete die Entdeckung des äußeren Fotoeffekts (↑ 1.2.2.2.) durch H. HERTZ im Jahre 1887. Dieser Effekt bleibt unverständlich, wenn man Licht ausschließlich als Welle betrachtet. Die von EINSTEIN 1905 angegebene Theorie des äußeren Fotoeffekts brachte mit der *Lichtquantenhypothese* eine Wiederbelebung der Teilchenauffassung. Weitere Beweise für die Richtigkeit der Lichtquantenhypothese brachte der 1923 entdeckte *Comptoneffekt* (↑ 1.2.2.3.) sowie die Erzeugung und Vernichtung von Elektron-Positron-Paaren (↑ 1.2.2.4.).

Das Wesen des Lichts wurde zu Beginn des 20. Jahrhunderts im Zusammenhang mit der Entwicklung der Quantenphysik erkannt. Seine Beschreibung erfordert die Benutzung beider Begriffe „Welle“ und „Teilchen“. Zur Bezeichnung dieses vom Standpunkt der klassischen Physik unverständlichen Verhaltens von Mikroobjekten wurde der Begriff Welle-Teilchen-Dualismus geprägt (↑ 1.2.4.).

1.2.2.2. Äußerer Fotoeffekt. Fällt elektromagnetische Strahlung auf einen Körper, so beobachtet man eine Streustrahlung unveränderter Wellenlänge. Sie entsteht durch einen als *Rayleighstreuung* bezeichneten Vorgang, der aufgrund der klassischen Wellenvorstellung vom Licht verständlich ist: Die einfallende Welle regt die Atomelektronen zu Schwingungen an, und diese emittieren als beschleunigt bewegte Ladungsträger Strahlung in alle Raumrichtungen, ausgenommen die Bewegungsrichtung.

Ist die Wellenlänge der Primärstrahlung kleiner als eine vom bestrahlten Material abhängige Grenze, so treten aus dem Körper zusätzlich Elektronen aus. Dieser als *äußerer Fotoeffekt* bezeichnete Prozeß wurde von HALLWACHS genauer untersucht und um 1900 durch LENARD quantitativ beschrieben.

Abb. 1.2.2-1 zeigt den grundsätzlichen Aufbau des Experiments. In einem evakuierten Glaskolben befindet sich eine durch ultraviolettes Licht bestrahlte Metallplatte, die als Fotoelektrode wirkt, und ein Gitter, um die befreiten Fotoelektronen aufzufangen. Zwischen beide ist eine regelbare Spannung U gelegt. Der jeweils fließende Fotostrom I kann am Meßgerät abgelesen werden. Daneben ist eine typische Kennlinie des Experiments mit Licht bestimmter Wellenlänge und Intensität dargestellt.

Bei positiver Gitterspannung entsteht der Sättigungsstrom I_{max}, indem alle Fotoelektronen durch das Gitter abgesaugt werden. I_{max} ist proportional zur Intensität des einfallenden Lichts. Bei negativer Gitterspannung gelangen Fotoelektronen nur nach Überwindung eines Bremspotentials zum Gitter. Der Abfall des Fotostroms zeigt, daß die Fotoelektronen

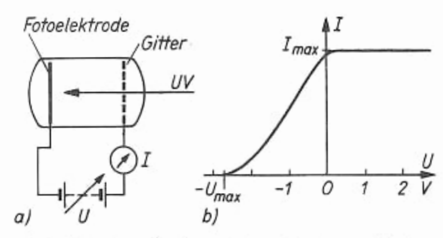


Abb. 1.2.2-1 Äußerer Fotoeffekt; a) Aufbau des Experiments, b) Strom-Spannungs-Kennlinie

unterschiedliche kinetische Energien besitzen, die aber jedenfalls nicht größer als eU_{max} sind. U_{max} hängt nicht von der Intensität des einfallenden Lichts ab, sondern wächst linear mit seiner Frequenz ν. Der Photonenstrom setzt unmittelbar mit der Bestrahlung ein. LAWRENCE und BEAMS fanden 1928, daß der zeitliche Abstand beider Ereignisse nicht größer als einige 10^{-9} s ist.
Vom Standpunkt der klassischen Physik führt ein Elektron unter dem Einfluß der einfallenden elektromagnetischen Welle erzwungene Schwingungen aus. Im Resonanzfall sollte es den Atomverband verlassen und als Fotoelektron ins Vakuum austreten können. Seine kinetische Energie müßte dann proportional zur Lichtintensität sein. Die Rechnung zeigt, daß es Tage benötigen würde, um auf die beschriebene Weise die beobachteten kinetischen Energien der Fotoelektronen aus der einfallenden Strahlung aufzusammeln. Diese Schlußfolgerungen widersprechen dem Experiment. Der äußere Fotoeffekt kann daher nicht aufgrund der klassischen Physik verstanden werden.
EINSTEIN erkannte das Wesen des Fotoeffekts: Energie wird von der Strahlung auf die Fotoelektronen nur in definierten Portionen $h\nu$ im Verlauf kurzzeitiger Stöße übertragen. Licht der Frequenz ν verhält sich dabei offenbar wie ein Strom von *Lichtquanten* oder *Photonen* der Energie $h\nu$, die sich mit Lichtgeschwindigkeit bewegen und daher entsprechend der Relativitätstheorie die Ruhmasse Null besitzen. Die Intensität des Lichts ist ein Maß für die Photonenstromdichte. Für die maximale Bewegungsenergie eines Fotoelektrons ergibt sich die Einsteinsche Beziehung $eU_{max} = h\nu - W_A$, wobei W_A eine materialspezifische Ablösearbeit von wenigen Elektronenvolt bedeutet.
Die Theorie von EINSTEIN beschreibt alle Beobachtungen richtig:
Die Befreiung des Fotoelektrons durch Stoß geschieht in mikroskopisch kleinen Zeiten, und die maximale Elektronenenergie ist proportional zur Frequenz sowie unabhängig von der Intensität des Lichts. Darüber hinaus muß die Zahl der je Zeiteinheit befreiten Fotoelektronen und damit I_{max} proportional zur Lichtintensität sein.
Das Revolutionäre an der Lichtquantenhypothese von EINSTEIN besteht darin, daß er der Planckschen Konstante h allgemeine Bedeutung (↑ 1.2.5.2.) zuerkannte. Wie er am Beispiel der Lichtquanten zeigte, stellt sie in der Beziehung $E = h\nu$ den Zusammenhang zwischen der Welleneigenschaft Frequenz und der Teilcheneigenschaft Photonenenergie her. Die Lichtquantenhypothese von EINSTEIN gewann an Überzeugungskraft, nachdem es MILLIKAN 1916 gelungen war, den Wert der Planckschen Konstante aus den maximalen Energien der Fotoelektronen bei verschiedenen Frequenzen der einfallenden Strahlung zu bestimmen. Der so ermittelte Wert wich nur um 0,5% von dem Wert ab, den PLANCK aus der Analyse der Hohlraumstrahlung (↑ 1.2.5.1.) erhalten hatte.

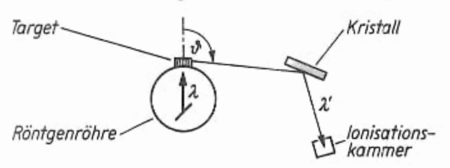


Abb. 1.2.2-2 Nachweis des Comptoneffekts

1.2.2.3. Comptoneffekt. Bei genügend kurzwelliger Primärstrahlung wird eine dritte Komponente der emittierten Strahlung intensiv, deren Wellenlänge größer als die Primärwellenlänge ist. Diese als *Comptoneffekt* bezeichnete Erscheinung kann anhand eines Experiments entsprechend Abb. 1.2.2-2 quantitativ untersucht werden: Röntgenstrahlung bekannter Wellenlänge λ fällt auf einen Körper. Aus der unter dem Winkel ϑ austretenden Strahlung wird durch Beugung an einem Kristall bekannter Struktur die Comptonstrahlung der Wellenlänge λ' entnommen und in der Ionisationskammer nachgewiesen. Es zeigt sich, daß die Wellenlängenverschiebung $\Delta\lambda = \lambda' - \lambda$ mit wachsendem Streuwinkel ϑ zunimmt. Sie ist unabhängig von der Wellenlänge und vom Material des bestrahlten Körpers.
Wegen der Materialunabhängigkeit des Comptoneffekts muß seine Ursache in der Wechselwirkung der Strahlung mit freien Elektronen gesucht werden. Ein Elektron gilt dabei als frei, wenn die Energie $h\nu$ des einwirkenden Photons groß im Vergleich mit der Energie seiner Bindung an den Körper ist. Die Interpretation des Comptoneffekts als Reflexion von Licht am bewegten Elektron im Sinne des *Dopplereffekts* der klassischen Physik liefert im Widerspruch zum Experiment $\Delta\lambda = f(\lambda)$ und scheidet damit aus. Die Quantentheorie des Comptoneffekts beschreibt diesen als einen durch Energie- und Impulserhaltung gekennzeichneten Zusammenstoß zwischen einem Photon und einem freien Elektron, dem *Rückstoßelektron*. Das Photon ist dabei gekennzeichnet durch Energie $h\nu$ und Impuls $p = h/\lambda$ mit $\lambda = c/\nu$ entsprechend der relativistischen Energie-Impuls-Beziehung $E = \sqrt{(mc^2)^2 + (cp)^2}$ für verschwindende Ruhmasse ($m = 0$). Energie- und Impulserhaltungssatz sind erfüllt, indem ein Photon kleinerer Frequenz in einer mit der Elektronenbewegung abgestimmten Richtung emittiert wird. Die Auswertung der Erhaltungssätze liefert in Übereinstimmung mit dem Experiment $\Delta\lambda = \Lambda_e(1 - \cos\vartheta)$. Dabei ist $\Lambda_e = h/mc = 2{,}426 \cdot 10^{-12}$ m mit der Elektronenruhmasse m

die sogenannte *Comptonwellenlänge* des Elektrons.
Die Theorie von COMPTON wurde durch eine Reihe weiterer Experimente überzeugend bestätigt. So fanden HOFSTADTER und MC INTYRE 1947, daß Rückstoßelektron und gestreutes Photon innerhalb von 10^{-9} s „gleichzeitig" erscheinen, nachdem BOTHE und GEIGER 1925 vorerst nur 10^{-3} s gemessen hatten. BLESS zeigte 1927, daß auch die Energieverteilung der Rückstoßelektronen mit der Rechnung von COMPTON übereinstimmt.
Der Comptoneffekt weist einerseits die quantenhafte Struktur der elektromagnetischen Strahlung nach und ist andererseits ein Beweis dafür, daß auch Prozesse zwischen Mikroobjekten durch Energie- und Impulserhaltungssatz geprägt sind.

1.2.2.4. Paarerzeugung und Paarvernichtung. Wird die Photonenenergie der einfallenden Strahlen größer als die doppelte Ruhenergie des Elektrons, $2mc^2 = 1{,}022$ MeV, so tritt zunehmend die als *Paarerzeugung* bezeichnete Umwandlung eines Photons in ein Elektron-Positron-Paar auf. Das *Positron*, das 1932 von ANDERSON in der kosmischen Strahlung entdeckt wurde, ist das *Antiteilchen des Elektrons* (↑ 1.1.6.1.). Die Paarerzeugung verläuft unter strenger Gültigkeit der Erhaltungssätze von Ladung, Energie und Impuls. Da Elektron und Positron nicht den gesamten Photonenimpuls aufnehmen können, findet der Prozeß unter Beteiligung eines weiteren Stoßpartners statt, meist ist das ein schwerer Kern.
Auch der Umkehrprozeß, die *Paarvernichtung*, wird beobachtet. Treffen Elektron und Positron aufeinander, so entstehen daraus zwei Photonen. Falls ein schwerer Kern zur Impulsaufnahme zur Verfügung steht, wird die Emission eines Photons der Energie $h\nu > 1{,}022$ MeV möglich. Paarerzeugung und Paarvernichtung verdeutlichen besonders eindringlich die Gültigkeit der Lichtquantenhypothese.

1.2.3. Materiewellen

1.2.3.1. De Broglie-Hypothese. Die Lichtquantenhypothese führte zu einer unbefriedigenden Asymmetrie des physikalischen Weltbilds: Licht konnte sich wie ein Teilchenstrom verhalten, während entsprechende Hinweise auf Welleneigenschaften von Elektronen zunächst fehlten. Andererseits verhielt sich auch das Elektron kaum wie ein klassisches Teilchen, wie die Bohrschen Postulate deutlich zeigten. Deshalb fand die 1924 von DE BROGLIE formulierte Hypothese der *Materiewellen* reges Interesse. In ihr wurde jedem physikalischen Objekt eine sogenannte *Phasenwelle* oder Materiewelle, wie sie später genannt wurde, zugeordnet. Die Materiewelle eines freien Teilchens ist im einfachsten Falle eine ebene Welle (↑ 2.3.1.2.). Wellenlänge und Frequenz hängen mit den Teilcheneigenschaften Impuls p und Energie E durch die Beziehungen $p = h/\lambda$ und $E = h\nu$ zusammen. Für elektromagnetische Wellen und Photonen waren diese Gleichungen damals schon experimentell bewiesen (↑ 1.2.2.2., 1.2.2.3.). DE BROGLIE forderte ein einheitliches Teilchen-Welle-Verhalten aller physikalischen Objekte auf der Basis dieser Gleichungen.
Durch geeignete Überlagerung ebener Materiewellen können beliebig konzentrierte Wellenfelder, sogenannte *Wellenpakete*, hergestellt werden. Diese enthalten den für Teilchen charakteristischen Aspekt der Lokalisierbarkeit und sind daher typische Materiewellen freier Teilchen. Die Bewegung eines Wellenpakets als Ganzes wird durch seine *Gruppengeschwindigkeit* (↑ 2.3.1.5.) beschrieben; sie stimmt mit der Geschwindigkeit des zugeordneten Objekts überein. Die Breite eines Wellenpakets aus Materiewellen nimmt ständig zu, da sich die überlagerten Teilwellen mit unterschiedlichen Geschwindigkeiten ausbreiten.

1.2.3.2. Wellenmechanik. Ausgehend von diesen Konsequenzen der Hypothese von DE BROGLIE entwickelte sich die *Wellenmechanik*. Sie führte nach Überwindung von Interpretationsschwierigkeiten zum quantitativen Verständnis atomarer Systeme.
Ursprünglich beschrieb die Wellenmechanik jedes Objekt als räumlich ausgedehnte Masseverteilung, deren Dichteverlauf durch die Intensität $|\psi(\boldsymbol{r}, t)|^2$ des jeweiligen Materiewellenfelds mit der *Wellenfunktion* $\psi(\boldsymbol{r}, t)$ als Funktion des Orts $\boldsymbol{r}$ und der Zeit t gegeben ist. Später (↑ 1.2.4.) wurde die Intensität als Dichte der Aufenthaltswahrscheinlichkeit des Objekts interpretiert. Zum Ausbau der Wellenmechanik trug vor allem SCHRÖDINGER bei, indem er 1926 das Problem der Berechnung der Wellenfunktion für ein Teilchen im äußeren Potential löste. Er verallgemeinerte die Wellengleichung der Materiewellen freier Teilchen zur sogenannten *Schrödingergleichung*, einer heute nach ihm benannten Grundgleichung der modernen Quantenmechanik. Im eindimensionalen Fall mit dem äußeren Potential $V(x)$ lautet sie:

$$-\frac{\hbar}{\mathrm{i}}\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V(x)\,\psi(x, t).$$

Dabei ist $\hbar \equiv h/2\pi$. Im Unterschied zur älteren Quantentheorie (↑ 1.1.4.6.) gelingt es der Wellenmechanik ohne Rückgriff auf klassische Teilchenvorstellungen vom Elektron, die Existenz der stationären Zustände der Atome und ihrer diskreten Energien zu begründen. Ähnlich den diskreten Schwingungsmoden einer Saite bilden sich in stationären Zuständen stehende Materiewellenfelder (s. a. 2.3.3.), deren charak-

teristische Frequenzen die Energien der Zustände bestimmen.
In Analogie zur Streuung von Lichtwellen an optischen Inhomogenitäten folgt aus der Schrödingergleichung, daß Materiewellen an Gebieten mit ortsveränderlichem Potential gestreut werden. Die Streuwellen interferieren und liefern so die für Beugungseffekte typischen Intensitätsverteilungen (↑ 1.2.1.). Also muß es z. B. Elektronenbeugung geben.

1.2.3.3. Elektronenbeugung. Angeregt durch die Hypothese von DE BROGLIE verwies im Juli 1925 ELSASSER auf eine Reihe von Arbeiten, in denen er Hinweise auf Beugungseffekte von Elektronen sah. Seine Mitteilung fand aber kaum Beachtung, so daß der Nachweis der Elektronenbeugung erst nach Veröffentlichung der Arbeitsergebnisse von DAVISSON und GERMER im Jahre 1927 allgemein zur Kenntnis genommen wurde.
Die eigentliche Entdeckung der Elektronenbeugung fand aber schon im April 1925 statt, und zwar aufgrund eines Unfalls im Labor von DAVISSON. Er untersuchte die Rückstreuung von Elektronen (50 bis 200 eV) durch eine polykristalline, also aus vielen ungeordneten, mikroskopisch kleinen Kristalliten bestehende Nickeloberfläche. Die Apparatur war in einem evakuierten Glaskolben untergebracht. Gemessen wurde die Winkelabhängigkeit des Stroms der ohne Energieverlust reflektierten Elektronen. Er fand ziemlich glatte Meßkurven, die er aufgrund der Teilchenvorstellung vom Elektron erklären konnte. Dann zerstörte die Explosion einer mit flüssiger Luft gefüllten Flasche den Rezipienten, und das hoch erhitzte Nickelpräparat überzog sich mit einer Oxidschicht. DAVISSON entfernte diese in einer neuen Apparatur durch sorgfältiges Ausheizen in Wasserstoff. Bei erneuter Durchführung des Streuexperiments zeigten die Meßkurven viele Maxima, die wie Beugungseffekte empfindlich von Energie und Einfallsrichtung der Elektronen abhingen. Während des Ausheizens waren nämlich aus vielen kleinen wenige große Kristallite entstanden, und nur etwa 10 davon bestrahlte er in seinem Experiment mit Elektronen. Nach der de-Broglieschen Beziehung $\lambda = h/p$ betrug die Elektronenwellenlänge ungefähr 10^{-10} m. Sie war damit von der Größenordnung der Gitterkonstante des Nickelkristalls, und es mußten deutliche Interferenzen entstehen.
Auf die Arbeiten von SCHRÖDINGER zur Wellenmechanik aufmerksam gemacht, begannen DAVISSON und GERMER 1926 mit der systematischen Untersuchung der Elektronenbeugung an definierten Kristallflächen von Nickel. 1927 führten sie den quantitativen Beweis, daß die beobachteten Intensitätsmaxima die Beugungsmaxima der de-Broglieschen Materiewellen der einfallenden Elektronen sind.
Das von DAVISSON begründete Verfahren der Beugung langsamer Elektronen an Kristalloberflächen ist zu einer wichtigen Methode für die Bestimmung von Oberflächenstrukturen ausgebaut worden (Tafel 31). Nach dem Vorbild der 1915 von DEBYE und SCHERRER entwickelten Variante der Röntgenstrukturuntersuchung auf der Basis eines Kristallpulvers erzeugte THOMSON ebenfalls 1927 Elektroneninterferenzen bei Durchstrahlung polykristalliner Metallfolien. An vielen Beispielen konnte so gezeigt werden, daß Elektronen mit der *de-Broglie-Wellenlänge* $\lambda = h/p$ die gleichen Interferenzfiguren erzeugen wie Röntgenstrahlen gleicher Wellenlänge. Beide Verfahren stellen ähnliche Möglichkeiten zur Strukturuntersuchung atomarer Systeme dar.

1.2.4. Welle-Teilchen-Dualismus

Wie das Verhalten von Licht und Elektronen zeigt, können diese *weder* als Welle *noch* als Teilchenstrom im Sinne der klassischen Physik bezeichnet werden. In ihrem Wesen sind vielmehr beide Seiten auf eine als *Welle-Teilchen-Dualismus* bezeichnete Art und Weise vereinigt. Das Verhalten von Mikroobjekten kann anhand des Doppelspaltexperiments (↑ 1.2.1.) verdeutlicht werden. Führt man das Experiment mit einem einzelnen Objekt aus, so wird dieses an irgendeinem Punkt des Auffangschirms als Ganzes nachgewiesen. Bei der Lokalisierung verhalten sich Elektron und Photon daher wie Teilchen. Das *Elektron* ist also *keine Materiewelle*, denn dann müßte es der Interferenzfigur entsprechend auf dem Schirm als kontinuierlich verschmiertes Objekt erscheinen. Folglich ist die anfangs von der Wellenmechanik gegebene Interpretation der Intensität des Materiewellenfelds falsch.
Um den Auftreffort des Mikroobjekts auf dem Schirm (↑ Abb. 1.2.1-1) einfach beschreiben zu können, soll dieser in schmale Streifen der Breite Δ. parallel zu den Spalten eingeteilt werden. Führt man das Experiment N-mal durch, so fallen im Streifen n gerade ΔN_n Objekte ein. Die relative Auftreffhäufigkeit ist damit $\Delta N_n/N$. Bei einer sehr großen Anzahl von Versuchen wird deutlich, daß die relative Häufigkeitsverteilung die berechnete Intensitätsverteilung des Materiewellenfelds widerspiegelt, es gilt $\Delta N_n/N \sim |\psi_n|^2$, wobei ψ_n die Wellenfunktion etwa in der Streifenmitte bedeutet. Für jedes der voneinander unabhängigen Experimente mit jeweils einem Objekt gibt aber die beobachtete relative Häufigkeit die Wahrscheinlichkeit Δw_n an, mit der das Objekt im Streifen n auf den Schirm trifft. Diese Wahrscheinlichkeit muß auch proportional zur Streifenbreite Δx sein, so daß man

$\Delta w_n = |\psi_n|^2 \Delta x$ erhält. Allgemein liefert die Intensität $|\psi(r, t)|^2$ der Wellenfunktion die Wahrscheinlichkeit $dw(r, t) = |\psi(r, t)|^2 dV$ dafür, daß das Mikroobjekt zur Zeit t im Volumenelement dV um den Ort r herum bei einer Ortsmessung angetroffen wird. Diese Interpretation des Materiewellenfelds gab 1926 BORN. Sie beseitigt die anfänglichen Interpretationsschwierigkeiten der Wellenmechanik, indem sie Wellen- und Teilcheneigenschaften der Mikroobjekte nebeneinander berücksichtigt, und bildet ein Fundament der Quantenphysik.

Der Welle-Teilchen-Dualismus kommt z. B. im Versuchsablauf beim Comptoneffekt zum Ausdruck (↑ Abb. 1.2.2-2): Die in der Ionisationskammer aufgrund ihrer Teilcheneigenschaften registrierte Strahlung zeigt bei der Beugung am Kristall ihre Welleneigenschaften und kurz zuvor bei der Wechselwirkung mit freien Elektronen wiederum ihre Teilcheneigenschaften. In jeder dieser drei Positionen des Experiments von COMPTON werden nicht gleichzeitig beide Aspekte der Strahlung beliebig genau sichtbar. Die Gestaltung einer experimentellen Anordnung zum Nachweis des Wellencharakters macht sie in gewissem Maße unbrauchbar, um den Teilchencharakter nachzuweisen, und umgekehrt. BOHR hat für ein derartiges Verhältnis zweier Eigenschaften oder Erscheinungsformen den Begriff *Komplementarität* geprägt.

Teilchen- und Welleneigenschaften der Mikroobjekte sind komplementär zueinander. Dabei werden ihre Ausbreitung durch die Wellen- und ihre Wechselwirkung mit atomaren Systemen durch die Teilcheneigenschaften gekennzeichnet.

Die zentralen Beziehungen des Welle-Teilchen-Dualismus sind die Einstein-de-Broglieschen Relationen $E = h\nu$ und $p = h/\lambda$. Sie stellen den Zusammenhang zwischen Teilchen- und Welleneigenschaften des Objekts her, wobei die *Plancksche Konstante* h als fundamentale Naturkonstante auftritt. Ihr außerordentlich kleiner Wert von $6{,}626 \cdot 10^{-34}$ Js hat zur Folge, daß die Wellenmechanik für makroskopische Körper keine Korrekturen zur klassischen Teilchenmechanik bringt.

Die Frage, ob Elektronen und Photonen Teilchen oder Wellen sind, erweist sich somit als Scheinfrage. Zur Beschreibung eines Mikroobjekts benötigt man beide Begriffe. Würde es sich beim Doppelspaltexperiment wie ein klassisches Teilchen auf einer Bahn durch einen bestimmten Spalt bewegen, so könnte es die Minima der Interferenzfigur auf dem Schirm nicht meiden, da es vom anderen Spalt überhaupt nicht beeinflußt wäre. Elektronen und Photonen bewegen sich daher nicht auf Bahnkurven; ihre Bewegung unterliegt Gesetzen der Wellenausbreitung.

1.2.5. Wirkungsquantum

1.2.5.1. Plancksche Entdeckung. Befinden sich die aus einer beliebigen physikalischen Substanz bestehenden Wände eines Hohlraums bei einer gegebenen Temperatur T im thermodynamischen Gleichgewicht, so ist dieser Hohlraum als Folge eines ständigen Wechselspiels von Emissions- und Absorptionsprozessen von elektromagnetischer Strahlung erfüllt, die sich in diesem Sinne ebenfalls im thermischen Gleichgewicht befindet. Diese Strahlung kann durch eine kleine, die Verhältnisse im Innenraum nur sehr wenig störende Öffnung in den Außenraum austreten und dort analysiert werden. Bereits 1860 erkannte KIRCHHOFF, daß diese Strahlung aufgrund der Gesetze der Thermodynamik sehr allgemeine, vom Wandmaterial völlig unabhängige Eigenschaften haben muß: Analysiert man die Hohlraumstrahlung nach Frequenzen bzw. Wellenlängen, so hängt (bei gegebenem Volumen) die Strahlungsenergie pro Frequenzintervall nur von der betrachteten Frequenz und der Temperatur ab. Durch Präzisionsmessungen, die LUMMER und PRINGSHEIM in Berlin an der damaligen Physikalisch-Technischen Reichsanstalt durchführten, wobei auch das Streben der Industrie zur Entwicklung effektiverer Strahlungsquellen solche Experimente motivierte, wurde die temperaturabhängige spektrale Verteilung der thermischen Gleichgewichtsstrahlung sehr genau bestimmt. Es zeigte sich bald, daß weder das damals etablierte klassische Wellenbild noch ein Teilchenmodell der Strahlung die experimentellen Ergebnisse erklären konnte. PLANCK versuchte, aufbauend auf seinem tiefen Verständnis der Thermodynamik, das Strahlungsgesetz theoretisch zu erfassen. Am 14. Dezember 1900 gab er in einem Vortrag vor der Berliner Akademie eine physikalische Deutung des etwa 8 Wochen vorher von ihm aufgestellten, die experimentellen Daten genau beschreibenden Strahlungsgesetzes. Dabei formulierte er die revolutionäre Hypothese, daß das Strahlungsfeld mit den Wänden des Hohlraums Energie nur in diskreten Energiequanten $\Delta E = h\nu$ austauschen kann; ν ist die jeweilige Frequenz und $h = 6{,}626 \cdot 10^{-34}$ Ws^2 eine Konstante von der Dimension Energie × Zeit. Eine solche Größe bezeichnet man als „Wirkung", h demzufolge als *Wirkungsquantum*. Mit dieser Quantenhypothese wurde ein diskretes, sprunghaftes Element in die Theorie eingeführt, das der klassischen Kontinuumsauffassung völlig fremd war.

Es ist bemerkenswert, daß das zu Ehren seines Entdeckers auch als *Plancksche Konstante* bezeichnete Wirkungsquantum bei der Analyse der sehr komplexen Situation der thermischen Strahlung entdeckt wurde. Andererseits ver-

deckte dieser Umstand zunächst die elementare und fundamentale Bedeutung des Wirkungsquantums. Es blieb auch offen, ob der quantisierte Energieaustausch zwischen Strahlungsfeld und Wandmaterial auf besondere Eigenschaften der Strahlung oder der die Wand aufbauenden Atome (wozu PLANCK neigte) oder der Wechselwirkung zwischen beiden beruhte.

1.2.5.2. Fundamentale Bedeutung des Wirkungsquantums. Die Bedeutung von h als fundamentale Naturkonstante, die gleichrangig neben der Vakuumlichtgeschwindigkeit steht, wurde erst allmählich deutlich. 1905 wandte EINSTEIN die Plancksche Beziehung bei der Deutung des lichtelektrischen Effekts (↑ 1.2.2.2.) erstmals auf einen Elementarprozeß an und vollzog mit seiner Lichtquantenhypothese die Quantelung des Strahlungsfelds selbst. Einen weiteren ähnlichen Schritt ging EINSTEIN 1912 mit der Erklärung des fotochemischen Grundgesetzes: Die Frequenz des Lichts muß mindestens so groß sein, daß die Energie $h\nu$ eines absorbierten Lichtquants die für die Einleitung einer chemischen Reaktion notwendige Veränderung im molekularen Bereich bewirken kann, beispielsweise die Fotodissoziation eines Moleküls. 1916 erkannte BOHR, daß das Wirkungsquantum grundlegend ist für den Atombau: Nur unter Benutzung von h kann die Stabilität der Atome und ihr diskretes Energiespektrum theoretisch erfaßt werden (↑ 1.1.4.5.). DE BROGLIE und SCHRÖDINGER verankerten 1924 und 1926 mit der Hypothese der Materiewellen (↑ 1.2.3.1.) und ihrem weiteren Ausbau zur Wellenmechanik (↑ 1.2.3.2.) die Plancksche Konstante noch fester in den Fundamenten der Naturbeschreibung. Sie demonstrierten die Durchgängigkeit des Welle-Teilchen-Dualismus (↑ 1.2.4.) und damit gleichzeitig die Universalität der beiden fundamentalen Beziehungen $E = h\nu$ und $p = h/\lambda$, in denen dieser Dualismus erfaßt wird. Auch aus moderner Sicht sind diese beiden Relationen diejenigen, in denen der fundamentale Charakter von h am klarsten wird. Das Wirkungsquantum verbindet die Teilchen- und Welleneigenschaften der Mikroobjekte miteinander und gestattet so die Formulierung ihres komplementären Wesens. Gleichzeitig wird dadurch die Anwendbarkeit sowohl des klassischen Teilchens wie die des klassischen Wellenbilds eingeschränkt. Insbesondere zeigte 1927 HEISENBERG mit der von ihm aufgestellten Unbestimmtheitsrelation $\Delta p \, \Delta x \gtrsim h$, daß man einem Teilchen nur innerhalb der durch diese Ungleichung bestimmten Grenzen einen bestimmten Impuls p und einen bestimmten Ort x gleichzeitig zuordnen kann. Diese Grenzen sind wegen des kleinen Zahlenwerts von h bei der Bewegung makroskopischer Teilchen stets weit unterhalb der Meßfehler und damit belanglos. Bei typischen Quantenerscheinungen sind dagegen die Produkte aus den charakteristischen Impuls- und Längenwerten gerade von der Dimension des Wirkungsquantums.

1.2.5.3. Wirkungsquantum und physikalische Größen. Bereits die Bohrsche Theorie der stationären Elektronenbahnen (↑ 1.1.4.5.) im Wasserstoffatom macht deutlich, daß sich erst mit Hilfe von h überhaupt für das Atom charakteristische Größen, wie Längen und Energien, bilden lassen. Dies bleibt auch in der modernen Quantentheorie unverändert gültig. Der *Bohrsche Radius* $a_0 = 4\pi\varepsilon_0\hbar^2/e^2m$ bestimmt die Größenordnung der räumlichen Abmessungen der Atome, und damit auch die einfacher Moleküle oder die der Gitterkonstanten kristalliner Festkörper bzw. die der mittleren Abstände benachbarter Moleküle in Flüssigkeiten. Die *atomare Energieeinheit* $e^2/4\pi\varepsilon_0a_0 = me^4/(4\pi\varepsilon_0\hbar)^2$ bestimmt die Größenordnung der in der Atomhülle auftretenden Energien, also etwa der Anregungs- oder Ionisierungsenergien. Beliebige Drehimpulskomponenten sind stets ganz- oder halbzahlige Vielfache von $\hbar$. Das *Bohrsche Magneton* $\mu_B = e\hbar/2m$ bestimmt die Größe der magnetischen Momente elektronischen Ursprungs, das *Kernmagneton* $\mu_K = e\hbar/2M_p$ mit der Protonenmasse M_p die Momente der Nukleonen und Atomkerne. Der magnetische Fluß in den sogenannten Flußschläuchen, die als magnetfelderfüllte Bereiche manche im Restvolumen feldfreie Supraleiter durchziehen, ist stets ein ganzzahliges Vielfaches des *Flußquantums* $2eh$, worauf moderne Präzisionsbestimmungen des Produkts aus Elementarladung und Wirkungsquantum beruhen. Im thermodynamischen Gleichgewicht tauscht ein System mit seiner Umgebung bei elementaren Stoßprozessen im Mittel Energien der Größenordnung kT aus (↑ 2.4.4.2.). Dies genügt für die Anregung von Schwingungsfreiheitsgraden oder die Erzeugung von Lichtquanten, wenn deren Frequenz ν nach der Planckschen Beziehung der Relation $h\nu \lesssim kT$ genügt. Wesentlich höhere Frequenzen als die durch $h\nu = kT$ festgelegte werden thermisch nicht angeregt. Diese Gesetzmäßigkeit findet sich bei der Hohlraumstrahlung in Form des *Wienschen Verschiebungsgesetzes*, wonach das Maximum in der Frequenzverteilung der Strahlungsenergie (↑ 1.2.5.1.) entsprechend $h\nu_{max} = 2{,}82\,kT$ bei einer sich linear mit der absoluten Temperatur verschiebenden Frequenz liegt. Blickt man in einen rotglühenden Ofen, so beträgt die Temperatur etwa 10^3 K und dem roten Spektralbereich entspricht eine Frequenz von $10^{14}\,s^{-1}$, woraus bereits die richtige Größenordnung der Planckschen Konstante abgeschätzt werden kann.

2. Grundlegende physikalische Begriffe und Gesetze

2.1. Bewegung von Teilchen

2.1.1. Naturwissenschaftliche Begriffe und Bezeichnungen

Um die Vielfalt der Naturerscheinungen exakt beschreiben und gedanklich erfassen zu können, sind genau erklärte und eindeutig bezeichnete Begriffe erforderlich. Ohne ein Verständnis dieser Begriffe bleibt der Sinn naturwissenschaftlicher Aussagen dunkel. Für die moderne Naturwissenschaft ergibt sich diese Forderung nach Präzision zwangsläufig aus ihrem quantitativen, messenden Charakter. Alle großen Fortschritte im qualitativen Verständnis, in der Aufdeckung immer tieferer Zusammenhänge, sind undenkbar ohne genaue quantitative Untersuchungen.

Die naturwissenschaftlichen Begriffe werden gebildet und präzisiert im Zusammenhang mit wirklichen oder auch nur gedachten Experimenten. Da die Möglichkeiten solcher Experimente vom Stand der bereits erreichten Naturerkenntnis abhängen, gilt dasselbe auch für die Begriffsbildung. Die naturwissenschaftlichen Begriffe sind nicht von vornherein gegeben, sie können nicht im voraus definiert werden, um dann als „Instrument" der Naturerkenntnis zu dienen; sie entstehen vielmehr und entwickeln sich weiter mit der Entwicklung der gesamten Naturwissenschaft. Dabei bleibt das Begriffssystem häufig während längerer Zeiten einer evolutionären Entwicklung ungeändert, während tiefgreifende, umwälzende Entdeckungen zu einer Neufestlegung des Inhalts alter oder zur Bildung neuer Begriffe zwingen. Solche Entdeckungen führen im alten Begriffssystem zu Widersprüchen, die durch die erneuerten Begriffe aufgelöst werden.

Zur Bezeichnung naturwissenschaftlicher Begriffe dienen häufig Wörter der normalen Umgangssprache, z. B. Masse, Kraft, Zeit, Energie, nämlich dann, wenn diese Begriffe Vorläufer oder Analoga außerhalb der Naturwissenschaft haben. Während diese Wörter außerhalb der Wissenschaft ihre jeweilige Bedeutung meist beibehalten, werden sie beim Übergang in das wissenschaftliche Vokabular nur in einer ganz bestimmten, eingeengten Bedeutung gebraucht. Sehr viele Wörter, wie Reflexion, Interferenz, Polarisation, entstammen dem Lateinischen als der Gelehrtensprache früherer Jahrhunderte. Viele dieser Bezeichnungen kehren, besonders für die technischen Anwendungen der Naturwissenschaft, in die Umgangssprache zurück und werden dann vor allem auf diese Anwendungen bezogen. Auch innerhalb der Wissenschaft werden Bezeichnungen oft über ihren ursprünglichen Sinn hinaus benutzt. So meint man mit „Spektrum" nicht nur das Farbband, zu dem weißes Licht mit Prismen oder anderen Geräten auseinandergezogen werden kann, sondern spricht auch von einem Geschwindigkeits- oder einem Energiespektrum.

2.1.2. Raum, Zeit, Bezugssystem

2.1.2.1. Längen und räumliche Koordinaten. Zur Analyse von Bewegungsvorgängen ist es zunächst erforderlich, die gegenseitige Lage von Körpern quantitativ zu erfassen. Dies geschieht am zweckmäßigsten dadurch, daß ein (starrer) Körper in an sich willkürlicher Weise vor allen anderen Körpern ausgezeichnet wird, indem man ihn als *Bezugskörper* wählt. Auf oder in diesem Bezugskörper soll ein *Bezugspunkt* O markiert werden. Dann kann die *Entfernung* l eines beliebigen Punkts P von O gemessen, d. h. durch die gewählte *Längeneinheit* ausgedrückt werden.

Legt man drei nicht in einer Ebene liegende *Bezugsrichtungen* fest, so kann die Lage eines beliebigen Punkts P folgendermaßen bestimmt werden: Man gibt an, welche Entfernungen l_1, l_2 und l_3 in diesen Richtungen bzw. den zu ihnen entgegengesetzten Richtungen vom Bezugspunkt O ausgehend zurückzulegen sind, um den Punkt P zu erreichen. Dabei kommt es auf die Reihenfolge nicht an. Charakterisiert man diejenigen Längen, die entgegengesetzt zur jeweiligen Bezugsrichtung zu durchlaufen sind, mit negativem Vorzeichen, so bilden $(\pm l_1, \pm l_2, \pm l_3)$ die nunmehr vorzeichenbehafteten *Koordinaten* von P in einem im allgemeinen schiefwinkligen *Koordinatensystem*. Aus Zweckmäßigkeitsgründen benutzt man meist ein rechtwinkliges, kartesisches Koordinatensystem mit den Koordinaten (x, y, z). Der gewählte Bezugskörper kann durch den Bezugspunkt O und die drei Bezugsrichtungen idealisiert werden. Der Bezugspunkt hat die Koordinaten $(0, 0, 0)$.

2.1.2.2. Zeit und Zeitkoordinate. Ein periodischer Vorgang definiert eine beliebig oft wie-

derholbare Dauer. Er kann als Uhr und seine Periode als *Zeiteinheit* aufgefaßt werden. Man hat immer genauere natürliche Uhren gefunden. Lange Zeit bildeten die Rotation der Erde um ihre Achse bzw. ihr sich periodisch wiederholender Umlauf um die Sonne die besten Uhren: Astronomische Beobachtungen dienten zur Festlegung der Zeiteinheit. In neuerer Zeit hat man auf der Grundlage atomarer bzw. molekularer Schwingungsvorgänge noch genauer gehende Atomuhren entwickelt.

Physikalisch definiert sind zunächst nur *Zeitdauern*. Man kann aber jedem einzelnen Moment eine Zeit zuordnen, wenn man einen *Zeitnullpunkt* definiert. Die Zeitdauer von ihm bis zum betrachteten Zeitmoment gilt dann als die diesem Moment zugeordnete Zeit. Obwohl die Zeitdauer immer eine positive Größe ist, unterscheidet man alle Zeiten, die früher als der Nullpunkt sind, als negativ von denen, die später als der Nullpunkt sind; genauso, wie negative räumliche Koordinaten trotz positiver Längen auftreten. Physikalische Gründe für die Wahl eines speziellen Zeitnullpunkts sind nicht bekannt.

2.1.2.3. Bezugssysteme. Eine vollständige Beschreibung der physikalischen Fakten oder Ereignisse besteht darin, daß für alle Raumpunkte und Zeitmomente (Zeitpunkte) die erforderlichen Angaben über den Aufenthalt von Teilchen, die Stärke und Richtung von Feldern u. a. gemacht werden. Diese *raumzeitliche Beschreibung* erfolgt also notwendigerweise immer in einem bestimmten Bezugssystem. Da Längen- und Zeiteinheit immer die gleichen sind, können sich verschiedene Bezugssysteme unterscheiden:

durch die Wahl von Bezugspunkt und/oder Bezugsrichtungen bei gleichem Bezugskörper

durch die Wahl von Zeitnullpunkt und/oder positiver Zeitrichtung

durch gegeneinander bewegte Bezugskörper.

Von besonderem Interesse und von besonderer Problematik ist die letzte Möglichkeit; ihre Analyse führt zur Relativitätstheorie.

2.1.2.4. Inertialsysteme und Relativitätsprinzip. Für die bloße Beschreibung der physikalischen Ereignisse sind alle Bezugssysteme prinzipiell gleichwertig. Das viel weitergehende Ziel der Physik ist aber, Naturgesetze aufzufinden und zu formulieren. Hierfür erweist sich eine bestimmte Klasse von Bezugssystemen als besonders geeignet, die *Inertialsysteme*. Das sind Bezugssysteme, in denen alle Orte, Richtungen und Zeiten physikalisch gleichwertig sind. Man sagt dann, der Raum sei homogen und isotrop, die Zeit homogen. Die Existenz solcher Systeme ist eine Erfahrungstatsache und keine Selbstverständlichkeit. Die Experimente zeigen, daß ein Bezugssystem, in dem die Sonne ruht, zumindest in ausgezeichneter Näherung ein Inertialsystem darstellt. Ein mit der Erde festverbundenes Bezugssystem ist vor allem wegen der Rotation der Erde um ihre eigene Achse kein Inertialsystem; jedoch kann man häufig auch von den dadurch bewirkten Effekten absehen. Für ein in einem Inertialsystem ruhendes kräftefreies Teilchen besteht kein Grund zur Bewegung in irgendeiner Richtung. Im Zusammenhang mit diesem *mechanischen Trägheitsgesetz* ist der Begriff des Inertialsystems [*inertia* lat., Trägheit] zuerst gebildet worden.

Die Erfahrung zeigt, daß alle gleichförmig, d. h. mit konstanter Relativgeschwindigkeit der Bezugskörper, gegeneinander bewegten Bezugssysteme völlig gleichwertig sind. Ist eins von ihnen ein Inertialsystem, so sind daher alle Inertialsysteme. Das *Relativitätsprinzip*, das EINSTEIN der speziellen Relativitätstheorie zugrunde legte, verschärft diese Aussage durch die Forderung, daß *alle* Naturgesetze in allen Inertialsystemen die gleiche mathematische Form haben. Ein gegenüber einem Inertialsystem beschleunigt bewegtes Bezugssystem ist mit Sicherheit kein Inertialsystem.

2.1.2.5. Positive Zeitrichtung. Bei vielen einfachen Systemen könnte jeder Prozeß ebensogut in umgekehrter zeitlicher Reihenfolge ablaufen. Ein rückwärts laufender Film zeigt dann keinerlei paradoxe, d. h. den Naturgesetzen widersprechenden Vorgänge. Bei solchen Prozessen besteht kein echter Unterschied zwischen „früher" und „später"; dementsprechend ist die Festlegung der positiven Zeitrichtung, also die Vorzeichenwahl der Zeitkoordinate t, genauso willkürlich wie die entsprechenden Richtungs- und Vorzeichenfestlegungen bei den räumlichen Koordinaten.

Ein echter Unterschied zwischen „früher" und „später", zwischen der Vergangenheit und Zukunft, der doch von der Erfahrung des menschlichen Lebens her sehr deutlich ist, entsteht erst durch physikalische Vorgänge, die nicht umkehrbar sind. Solche *irreversiblen Prozesse* laufen ab, wenn ein makroskopisches System in seinen thermodynamischen Gleichgewichtszustand übergeht, z. B. durch Diffusion, Wärmeleitung oder innere Reibung. Eine weitere Auszeichnung einer Zeitrichtung bzw. Prozeßrichtung ergibt sich bei Strahlungsvorgängen, wo immer nur von der Strahlungsquelle auslaufende Wellen auftreten, aber keine einlaufenden. Schließlich hat durch die Expansion des Universums die Nichtgleichwertigkeit der beiden Zeitrichtungen auch einen wesentlichen kosmologischen Aspekt. Üblicherweise wird die Zeitachse so gewählt, daß bei zwei Ereignissen *1* und *2* mit den Zeitkoordinaten $t_1 < t_2$ das Ereignis *2* später als *1* stattfindet.

2.1.2.6. Skalare und Vektoren. Viele physikalische Größen können (nach Wahl der Maßeinheit) durch eine einzige Zahlenangabe bereits

vollständig charakterisiert werden, so z. B. Länge, Zeitdauer, Masse, Ladung, Druck, Temperatur. Solche Größen heißen *Skalare*. Zu ihrer eindeutigen Beschreibung sind die zum Bezugssystem gehörenden Bezugsachsen nicht notwendig, da diese Größen selbst keinen gerichteten Charakter haben. Beim Vergleich von gegeneinander bewegten Bezugssystemen ändern im allgemeinen alle physikalischen Größen ihre Werte; von den oben aufgeführten haben nur Ladung und Ruhmasse einen von der Wahl des Bezugssystems unabhängigen Wert.

Andere physikalische Größen haben gerichteten Charakter. Ihre vollständige Beschreibung erfordert, ihre „Stärke" und ihre Richtung anzugeben, so z. B. Kraft, elektrische Feldstärke, Verschiebung im Raum. Solche Größen heißen *Vektoren*. Sie können geometrisch durch einen Pfeil veranschaulicht werden, dessen Länge der „Stärke" der Größe entspricht. Diese „Stärke" ist eine skalare Größe und heißt *Betrag* des Vektors. Der Betrag hat die gleiche (physikalische) Dimension wie die Größe selbst. Aus der Gleichheit zweier Vektoren folgt, daß sowohl ihre Beträge als auch ihre Richtungen übereinstimmen. Vektoren werden durch halbfette Buchstaben $\boldsymbol{a}$ gekennzeichnet, ihre Beträge als $a = |\boldsymbol{a}|$.

Da eine Richtung nur relativ zu anderen Richtungen definiert werden kann, erfordert das Arbeiten mit Vektoren, zumindest bei zeitabhängigen Vektoren, die Festlegung von Bezugsachsen. Anstatt durch Betrag und Richtung kann ein Vektor auch durch seine Projektionen auf die Richtungen eines zweckmäßigerweise *rechtwinkligen Basissystems* (Bezugsachsen) eindeutig bestimmt werden. Diese Projektionen heißen *Komponenten*. Aus der Gleichheit zweier Vektoren folgt die Gleichheit ihrer Komponenten. Die Lage eines beliebigen Punkts P bezüglich eines Bezugskörpers kann durch seine Entfernung vom Bezugspunkt O und durch die Richtung, in der man von O auszugehen hat, eindeutig bestimmt werden. Von O zu P führt ein bestimmter Pfeil. Auf diese Weise ist jedem Raumpunkt ein *Ortsvektor* $\boldsymbol{r}$ zugeordnet, dessen Betrag $r = |\boldsymbol{r}|$ die Entfernung von O angibt.

2.1.2.7. Axiale Vektoren und Schraubenregel. Diejenigen Vektoren, deren Richtung wie in den obigen Beispielen unmittelbar gegeben ist, werden genauer als *polare Vektoren* bezeichnet. Daneben benutzt man *axiale Vektoren* oder *Pseudovektoren*. Sie entsprechen Größen vom geometrischen Charakter einer *Achse mit Drehsinn*, z. B. Winkelgeschwindigkeit oder magnetische Feldstärke. Einer solchen Drehachse kann nur mit Hilfe einer zusätzlichen willkürlichen Festsetzung eine Richtung, ein Pfeil, zugeordnet werden. Die übliche Konvention ist die *Rechtsschraubenregel*, nach der die Pfeilspitze in die Richtung zeigt, in die sich eine Rechtsschraube beim Drehen bewegt. Dreht man ein kartesisches Koordinatensystem so um die z-Achse, daß die positive x-Achse auf die positive y-Achse zugedreht wird und orientiert die z-Achse entsprechend dieser Rechtsschraubenregel, so bilden die Koordinatenachsen in

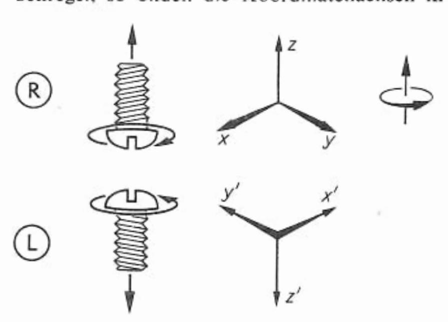


Abb. 2.1.2-1 Rechts- und Linksschraube, Rechts- und Linkssystem, Drehachse und axialer Vektor

der Reihenfolge x, y, z oder y, z, x oder z, x, y ein Rechtssystem (Abb. 2.1.2-1). In einem Linkssystem ist die z-Achse entgegengesetzt gerichtet, bzw. so, wie sich eine Linksschraube beim Eindrehen vorwärtsbewegt. Die Rechtsschraubenregel dient auch zur Festlegung der Richtung des vektoriellen Produkts zweier Vektoren. Das Vektorprodukt aus zwei polaren Vektoren ist also ein axialer Vektor. Wichtige Beispiele dafür sind *Drehimpuls und Drehmoment*. Dagegen ist das Vektorprodukt aus einem polaren und einem axialen Vektor wieder polar; dies trifft für die *magnetische Kraft* $q\boldsymbol{v} \times \boldsymbol{B}$ auf eine mit der Geschwindkeit $\boldsymbol{v}$ im Magnetfeld $\boldsymbol{B}$ bewegte Ladung q zu.

Während der Unterschied zwischen rechts und links anschaulich unmittelbar gegeben ist, ist die Zuordnung der Begriffe rechts und links zu den geometrisch verschiedenen Gebilden nur möglich, wenn man ein Standardobjekt mit der erforderlichen Asymmetrie als „rechts" oder „links" definiert. Die Benutzung axialer Vektoren setzt also einen solchen Standard voraus.

2.1.3. Impuls und Energie, Drehimpuls

2.1.3.1. Freie Teilchen. Als *Teilchen* soll ein Objekt bezeichnet werden, das soweit stabil und räumlich lokalisiert ist, daß man von seiner Bewegung als Ganzem sprechen kann. Das bedeutet vor allem, daß seine momentane Lage im Raum hinreichend genau durch einen einzigen Ortsvektor $\boldsymbol{r}$ charakterisiert ist, und daß das Teilchen eine definierte *Geschwindig-*

keit $\boldsymbol{v}$ besitzt. Der Betrag des Geschwindigkeitsvektors gibt an, wie schnell sich das Teilchen bewegt; seine Richtung, wohin es fliegt. Je nach der Aufgabenstellung kann z. B. ein Himmelskörper, eine Billardkugel, ein Atom oder ein Elektron als Teilchen betrachtet werden.

Die Geschwindigkeit hängt offensichtlich vom Bezugssystem ab. Geht man zu einem Bezugssystem über, das sich mit der Relativgeschwindigkeit $\boldsymbol{u}$ gegenüber dem ursprünglichen Bezugssystem bewegt, so ist $\boldsymbol{v} = \boldsymbol{V} - \boldsymbol{u}$ die *Teilchengeschwindigkeit im neuen Bezugssystem*, wenn sich das Teilchen im ursprünglichen Bezugssystem mit $\boldsymbol{V}$ bewegt. Dieser einfache Zusammenhang gilt nur bei Vernachlässigung relativistischer Effekte, d. h., wenn alle drei Geschwindigkeiten klein gegenüber der Lichtgeschwindigkeit c sind.

Ein Teilchen, das keinerlei äußeren Einwirkungen unterliegt, heißt *kräftefrei* oder kurz frei. In einem Inertialsystem sind alle Raumpunkte und alle Richtungen gleichwertig (↑ 2.1.2.4.). Ein in einem solchen System ruhendes freies Teilchen bleibt daher immer in Ruhe, führt aber in einem anderen Inertialsystem eine Bewegung mit konstanter Geschwindigkeit aus. Dies ist demnach der allgemeine Bewegungszustand eines kräftefreien Teilchens.

2.1.3.2. Stoß zweier gleicher Teilchen. Als Stoß soll ein Prozeß bezeichnet werden, bei dem zwei Teilchen mit Geschwindigkeiten $\boldsymbol{v}_1$ und $\boldsymbol{v}_2$ derart aufeinander zufliegen, daß sie sich in einem gewissen Zeitintervall soweit nähern, um miteinander in Wechselwirkung zu treten, sich also gegenseitig beeinflussen können (Abb. 2.1.3-1). Als Ergebnis dieser Wechselwirkung ändern beide Teilchen ihre Geschwindigkeit. Die Geschwindigkeiten nach dem

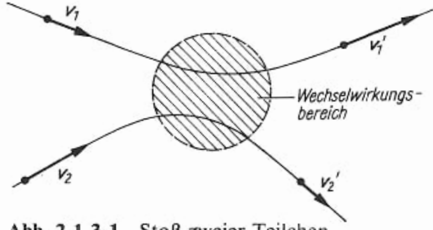


Abb. 2.1.3-1 Stoß zweier Teilchen

Stoß seien $\boldsymbol{v}_1'$ und $\boldsymbol{v}_2'$. Zwischen den Geschwindigkeiten $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ vor und den Geschwindigkeiten $\boldsymbol{v}_1'$, $\boldsymbol{v}_2'$ nach dem Stoß bestehen Zusammenhänge, die mit den Begriffen (kinetische) Energie und Impuls quantitativ erfaßt werden können.

Besonders einfach sind die Verhältnisse beim Stoß zweier gleicher Teilchen. Trotz dieser Vereinfachung sind im allgemeinen beide Partner beim Stoß- bzw. Wechselwirkungsprozeß nicht gleichwertig. Beispielsweise wird beim unmittelbaren Stoß zweier Billardkugeln die eine „vorn" sein und die andere sie „von hinten" anstoßen. Die Begriffe „vorn" und „hinten" beziehen sich dabei auf die Vorzugsrichtung der Bewegung beider Partner. Solche Unsymmetrien verschwinden nur dann völlig, wenn es eine solche Vorzugsrichtung nicht gibt, d. h., wenn beide Stoßpartner mit gleich großen und entgegengesetzt gerichteten Geschwindigkeiten, also $\boldsymbol{V}_1 = -\boldsymbol{V}_2$ oder $\boldsymbol{V}_1 + \boldsymbol{V}_2 = \boldsymbol{0}$, aufeinander zufliegen. Der Stoß braucht kein „zentraler Stoß" zu sein.

Die Teilchen selbst sollen durch den Wechselwirkungsprozeß nicht verändert werden, so daß nach dem Stoß exakt (auch bezüglich ihres „inneren Zustands") die gleichen Teilchen vorliegen wie vorher. Ein solcher Prozeß heißt *elastischer Stoß*. Dabei müssen im betrachteten Spezialfall:

erstens die Geschwindigkeiten beider Teilchen auch nach dem Stoß entgegengesetzt gleich groß sein, da anderenfalls die Vektorsumme $\boldsymbol{V}_1' + \boldsymbol{V}_2'$ ihrer Geschwindigkeiten eine bestimmte Raumrichtung ungerechtfertigterweise auszeichnen würde (↑ 2.1.2.4.);

zweitens ihre Geschwindigkeitsbeträge $V_1' = V_2' = V_1 = V_2$ genauso groß wie vor dem Stoß sein, da anderenfalls ein wesentlicher Unterschied zwischen dem ursprünglichen Stoßprozeß S und seinem Umkehrprozeß U bestände.

Der bisher untersuchte spezielle symmetrische Stoßprozeß erscheint in einem mit der konstanten Geschwindigkeit $\boldsymbol{u}$ bewegten Bezugssystem als ein allgemeiner, bezüglich der Geschwindigkeiten beider Teilchen unsymmetrischer Prozeß, bei dem $\boldsymbol{v}_1 + \boldsymbol{v}_2 \neq \boldsymbol{0}$ und $\boldsymbol{v}_1' + \boldsymbol{v}_2' \neq \boldsymbol{0}$ sind. Aus dem Relativitätsprinzip folgt, daß der Stoßprozeß im ursprünglichen Bezugssystem genauso abläuft, wenn man die Teilchen mit den Geschwindigkeiten $\boldsymbol{v}_1$ und $\boldsymbol{v}_2$ „aufeinanderschießt". Dabei werden mit $\boldsymbol{v}_1, \boldsymbol{v}_2$ bzw. $\boldsymbol{v}_1', \boldsymbol{v}_2'$ die Geschwindigkeiten im neuen Bezugssystem bzw. beim allgemeinen Stoß zweier gleicher Teilchen bezeichnet. Die Relationen zwischen diesen Geschwindigkeiten ergeben sich aus dem nichtrelativistischen Transformationsgesetz (↑ 2.1.3.1.) der Geschwindigkeiten. Es gilt:

$$\boldsymbol{v}_1 + \boldsymbol{v}_2 = (\boldsymbol{V}_1 - \boldsymbol{u}) + (\boldsymbol{V}_2 - \boldsymbol{u}) = -2\boldsymbol{u}$$

$$\boldsymbol{v}_1' + \boldsymbol{v}_2' = (\boldsymbol{V}_1' - \boldsymbol{u}) + (\boldsymbol{V}_2' - \boldsymbol{u}) = -2\boldsymbol{u},$$

wobei $\boldsymbol{V}_1 + \boldsymbol{V}_2 = \boldsymbol{0}$ und $\boldsymbol{V}_1' + \boldsymbol{V}_2' = \boldsymbol{0}$ beachtet wurde. Der Vergleich zeigt die Gültigkeit der Relation

$$\boldsymbol{v}_1 + \boldsymbol{v}_2 = \boldsymbol{v}_1' + \boldsymbol{v}_2'.$$

Es sind demnach zwischen zwei gleichen Teilchen nur solche Stöße möglich, bei denen sich

die *vektorielle Summe der Teilchengeschwindigkeiten* nicht ändert.

Eine zweite Relation ergibt sich, wenn man das Verhalten der Geschwindigkeitsquadrate untersucht. Das Quadrat $\boldsymbol{a}^2$ eines Vektors ist als Skalarprodukt $\boldsymbol{aa}$ des Vektors mit sich selbst definiert und stimmt mit dem Betragsquadrat a^2 überein. Eine einfache Rechnung zeigt, daß aus der Gleichheit aller Geschwindigkeitsquadrate für den symmetrischen Stoßprozeß die Relation $v_1^2 + v_2^2 = v_1'^2 + v_2'^2$ folgt. Auch bei einem allgemeinen elastischen Stoß zwischen zwei gleichen Teilchen bleibt also die *Summe der Geschwindigkeitsquadrate* konstant, obwohl sich die Geschwindigkeiten selbst im allgemeinen ändern. Das einfachste *Beispiel* ist der zentrale Stoß eines Teilchens auf ein vor dem Stoß ruhendes zweites (gleiches) Teilchen. Das vorher ruhende Teilchen fliegt nach dem Stoß mit der Geschwindigkeit des stoßenden Teilchens fort, während dieses liegen bleibt.

2.1.3.3. Masse. Beim Stoß zweier gleicher Teilchen unterscheiden sich entsprechend $\boldsymbol{V}_1' - \boldsymbol{V}_1 = -(\boldsymbol{V}_2' - \boldsymbol{V}_2)$ die vektoriellen Geschwindigkeitsänderungen beider Teilchen nur durch das Vorzeichen. Beim Stoß zweier verschiedener Teilchen ist das nicht mehr so: Das „schwerere" Teilchen ist „träger" und ändert seine Geschwindigkeit weniger als das „leichtere". Dieser Effekt wird berücksichtigt, indem man jedes Teilchen durch seine *Masse m*, das ist ein seine *Trägheit* charakterisierender Parameter, charakterisiert. Die Masse ist eine wichtige Eigenschaft, ein Merkmal des Teilchens. Bei verschiedenen Wechselwirkungen, bespielsweise bei Stößen mit verschiedenen Partnern, zeigt ein Teilchen immer die gleiche Masse.

Die Masse kann aufgrund folgender Erfahrungen exakt definiert und gemessen werden: Man schießt zwei Teilchen so aufeinander, daß $\boldsymbol{V}_1$ und $\boldsymbol{V}_2$ antiparallel sind und variiert das Geschwindigkeitsverhältnis V_1/V_2 solange, bis auch die Geschwindigkeiten $\boldsymbol{V}_1'$ und $\boldsymbol{V}_2'$ nach dem Stoß immer (d. h. bei beliebig „schiefem Stoß") antiparallel sind. Das Experiment zeigt, daß dies möglich ist, daß dann keines der beiden Teilchen den Betrag seiner Geschwindigkeit ändert, also $V_1 = V_1'$ und $V_2 = V_2'$, und daß dann das Geschwindigkeitsverhältnis beider Teilchen (unabhängig von V_1) eine für das stoßende Teilchenpaar charakteristische Konstante ist, also $V_1/V_2 = V_1'/V_2' = \alpha_{12}$. Definiert man durch $\alpha_{12} \equiv m_1/m_2$ das *Verhältnis der Massen* beider Teilchen, so gilt für die ausgewählte spezielle Klasse von Stoßprozessen

$$(1) \qquad m_1\boldsymbol{V}_1 + m_2\boldsymbol{V}_2 = m_1\boldsymbol{V}_1' + m_2\boldsymbol{V}_2' = \boldsymbol{0}.$$

Die Geschwindigkeitsänderungen $\boldsymbol{V}_1' - \boldsymbol{V}_1$ und $\boldsymbol{V}_2' - \boldsymbol{V}_2$ verhalten sich zueinander umgekehrt wie die Teilchenmassen.

Bestimmt man zunächst die Massen m_1 und m_2 zweier Teilchen und danach die Masse m_{12} des aus beiden durch Verschmelzung entstehenden Teilchens, so zeigt das Experiment, daß $m_{12} \approx m_1 + m_2$ gilt, d. h. die *Additivität der Massen* (s. a. 2.1.3.6.).

2.1.3.4. Impuls und kinetische Energie. Von der Beziehung 2.1.(1) ausgehend können alle für den Spezialfall $m_1 = m_2$ bereits durchgeführten Überlegungen (↑ 2.1.3.2.) fast unverändert übernommen werden. Die Relation $\boldsymbol{v}_1 + \boldsymbol{v}_2 = \boldsymbol{v}_1' + \boldsymbol{v}_2'$ verallgemeinernd gilt

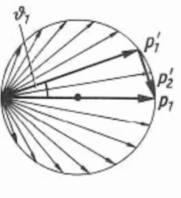
$$(2) \qquad m_1\boldsymbol{v}_1 + m_2\boldsymbol{v}_2 = m_1\boldsymbol{v}_1' + m_2\boldsymbol{v}_2' \quad \text{bzw.} \quad \boldsymbol{p}_1 + \boldsymbol{p}_2 = \boldsymbol{p}_1' + \boldsymbol{p}_2'$$

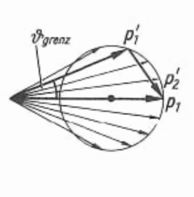
für *jeden* Stoß zweier Teilchen mit den Massen m_1 und m_2. Die vektorielle Größe $\boldsymbol{p} \equiv m\boldsymbol{v}$ heißt *Impuls* des Teilchens; der Impulsvektor hat die gleiche Richtung wie die Geschwindigkeit. Die obigen Gleichungen formulieren den kurz als Impulssatz bezeichneten *Impulserhaltungssatz*: Die Vektorsumme der Teilchenimpulse wird durch die Wechselwirkung zwischen den Teilchen nicht verändert und bleibt daher beim Stoß konstant. Darin besteht die Bedeutung des Impulsbegriffs.

Die Relation $v_1^2 + v_2^2 = v_1'^2 + v_2'^2$ verallgemeinernd gilt

$$(3) \qquad \frac{m_1}{2} v_1^2 + \frac{m_2}{2} v_2^2 = \frac{m_1}{2} v_1'^2 + \frac{m_2}{2} v_2'^2 \quad \text{bzw.} \quad \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} = \frac{p_1'^2}{2m_1} + \frac{p_2'^2}{2m_2}.$$

Die skalare Größe $E = mv^2/2 = p^2/2m$ heißt *kinetische Energie* des Teilchens. Der hier willkürlich erscheinende Faktor 1/2 ergibt sich



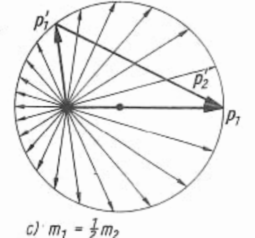


Abb. 2.1.3-2 Erlaubte Impulse nach dem Stoß eines Teilchens der Masse m_1 auf ein vorher ruhendes Teilchen der Masse m_2 für verschiedene Massenverhältnisse

zwanglos, wenn man von dem relativistisch exakten Energieausdruck (↑ 2.1.3.5.) ausgeht; er sichert außerdem einen einfachen Zusammenhang zwischen kinetischer Energie und Arbeit. Die obigen Gleichungen formulieren den kurz als Energiesatz bezeichneten *Energieerhaltungssatz*: Die Summe der kinetischen Energien bleibt beim Stoß unverändert. Darin besteht die Bedeutung dieses Energiebegriffs.

Energie- und Impulssatz stellen eine wesentliche Einschränkung für die bei der Wechselwirkung zweier Teilchen auftretenden Prozesse dar. Die nach dem Stoß möglichen Impulse hängen vom Massenverhältnis beider Teilchen ab (Abb. 2.1.3-2). Die graphische Konstruktion geht von der Teilung des „Primärimpulses" $\boldsymbol{p}_1$ im Verhältnis der Massen aus. $\boldsymbol{p}_1'$ kann irgendeiner der auf der Kugel endenden Vektoren sein. $\boldsymbol{p}_2'$ ergibt sich dann aus $\boldsymbol{p}_1' + \boldsymbol{p}_2' = \boldsymbol{p}_1$. Es ist ϑ_1 der Winkel, um den das stoßende Teilchen abgelenkt wird. Für $m_1 = m_2$ kann $\boldsymbol{p}_1' = \boldsymbol{0}$ sein: Beim zentralen Stoß gleicher Teilchen erfolgt ein *vollständiger Energieaustausch*. Bei $m_1 > m_2$ existiert ein Grenzwinkel ϑ_{grenz} für die Ablenkung des stoßenden Teilchens, während dieses bei $m_1 < m_2$, also beim Stoß mit einem schwereren Teilchen, auch „zurückgestreut" werden kann. Ein auf ein viel schwereres ruhendes Teilchen stoßendes Teilchen wird schließlich von diesem praktisch ohne Energieänderung „reflektiert".

2.1.3.5. Schnell bewegte Teilchen. Die für zwei massegleiche Teilchen (↑ 2.1.3.2.) durchgeführte Begründung von Impuls- und Energiesatz beruht auf dem nur für kleine Teilchengeschwindigkeiten gültigen Transformationsgesetz $\boldsymbol{v} = \boldsymbol{V} - \boldsymbol{u}$ der Geschwindigkeiten. Im allgemeinen Fall sind die komplizierteren Transformationsgesetze der speziellen Relativitätstheorie zu benutzen. Sie führen zwingend zu veränderten Geschwindigkeitsabhängigkeiten von Impuls und Energie.

Der *relativistische Impulsausdruck* lautet

$$(4) \quad p = \frac{mv}{\sqrt{1 - v^2/c^2}} \equiv m(v)\, v \approx mv \quad \text{für } v \ll c.$$

Für $v \ll c$, d. h. für Teilchengeschwindigkeiten, die klein gegenüber der Lichtgeschwindigkeit c sind, vereinfacht sich der Ausdruck zu dem einfacheren $m\boldsymbol{v}$ der nichtrelativistischen Mechanik. Häufig wird $m(v) \equiv m/\sqrt{1 - v^2/c^2}$ als *geschwindigkeitsabhängige Masse* eingeführt. Dadurch erhält der Impuls $p = m(v)\, \boldsymbol{v}$ formal wieder die alte Form. Die Teilchenmasse m wird zur Vermeidung von Mißverständnissen *Ruhmasse* genannt und oft mit dem Symbol m_0 bezeichnet. Nähert sich v der Lichtgeschwindigkeit, so nimmt der Impuls beliebig große Werte an, ebenso die geschwindigkeitsabhängige Masse $m(v)$ (Abb. 2.1.3-3a). Im Bereich $v \lesssim c$ ist es offenbar viel zweckmäßiger, den Impuls anstelle der Geschwindigkeit zur Charakterisierung des Bewegungszustands zu benutzen. Die Geschwindigkeit eines Teilchens mit von Null verschiedener Ruhmasse ist immer kleiner als c.

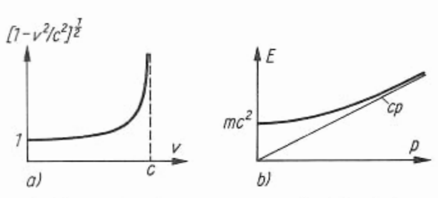


Abb. 2.1.3-3 a) Zur geschwindigkeitsabhängigen Masse; b) Zusammenhang zwischen Energie und Impuls

Der *relativistische Energieausdruck* lautet

$$(5) \quad E = \frac{mc^2}{\sqrt{1 - v^2/c^2}} = m(v)\, c^2 \approx mc^2 + \frac{m}{2} v^2 \quad \text{für} \quad v \ll c.$$

Dieser Energieausdruck nimmt für $v = 0$ den Wert mc^2 an, der als *Ruhenergie* bezeichnet wird. Solange die Teilchen und damit ihre Ruhmassen beim Stoß unverändert bleiben (s. a. 2.1.3.6.), ist mc^2 eine unwichtige Konstante, d. h., man könnte auch $E - mc^2$ als Energie des Teilchens definieren. Tatsächlich geht $E - mc^2$ für kleine Geschwindigkeiten $v \ll c$ in den nichtrelativistischen Ausdruck $mv^2/2$ für die kinetische Energie über. Nähert sich v der Lichtgeschwindigkeit, so nimmt E ebenso wie p beliebig große Werte an.

Besonders bei großen Geschwindigkeiten $v \approx c$ ist es zweckmäßig, die Geschwindigkeit v zu eliminieren und nur den Zusammenhang zwischen E und p zu betrachten. Eine einfache Rechnung liefert

$$(6) \quad E = c\sqrt{(mc)^2 + p^2} \approx mc^2 + \frac{p^2}{2m} \quad \text{für } p \ll mc \quad \text{bzw.} \quad v \ll c.$$

Für kleine Geschwindigkeiten erscheint wieder die einfache Form $p^2/2m$ der kinetischen Energie. Für sehr große Impulse wird schließlich $p \gg mc$ und $E \approx cp$ (Abb. 2.1.3-3b). Teilchen mit einem solchen genähert linearen Energie-Impuls-Zusammenhang werden oft *ultrarelativistisch* genannt. Die Geschwindigkeit kann in der Form

$$(7) \quad v = \frac{cp}{\sqrt{(mc^2) + p^2}} < c$$

durch den Impuls ausgedrückt werden. Es gilt $v = \mathrm{d}E(p)/\mathrm{d}p$.

Aus den letzten beiden Gleichungen für $E(p)$ und $v(p)$ wird deutlich, daß es Teilchen ohne Ruhmasse, also mit $m = 0$, geben kann. Sie bewegen sich immer mit Lichtgeschwindigkeit, $v = c$, und sind durch den linearen Zusammenhang $E = cp$ zwischen Energie und Impuls charakterisiert. Photonen sind solche Teilchen.

2.1.3.6. Allgemeine Stoßprozesse, Ruhenergie und Masse. Die bisher betrachteten elastischen Stöße sind dadurch definiert, daß vor und nach dem Stoß exakt dieselben Teilchen vorliegen. Bei allgemeineren Wechselwirkungsprozessen ist diese Voraussetzung im allgemeinen nicht mehr erfüllt. Bei der Wechselwirkung können „neue" Teilchen entstehen, wobei auch deren Anzahl nicht festgelegt ist. So können zwei Teilchen miteinander zu einem „verschmelzen" oder umgekehrt durch „Zerfall" eines Teilchens entstehen. Ein instabiles Teilchen kann in mehr als zwei andere zerfallen, oder beim Stoß zweier Teilchen können drei oder mehr Teilchen gebildet werden. Insbesondere können nach dem Stoß zwei Teilchen mit anderen Massen als vorher auftreten.

Wechselwirkungsprozesse allgemeinen Typs sind zwar für die Entstehung, Vernichtung und Umwandlung von Elementarteilchen besonders typisch, aber keineswegs auf diese beschränkt. Schießt man ein Projektil in einen Sandsack, so wird es von diesem „absorbiert"; ein Teilchen „verschwindet". Reißt der Faden, der zunächst das Entspannen einer zwischen zwei Körpern zusammengedrückten Feder verhindert, so „zerfällt" das ursprüngliche instabile Teilchen in zwei „Bruchstücke".

Für alle Prozesse der geschilderten Art gelten ohne Einschränkung die *Erhaltungssätze für Impuls und Energie* in der Form, daß die Vektorsumme aller Teilchenimpulse und die Summe aller Teilchenenergien vor und nach dem Stoß gleich sind, also

$$\sum_i^{(\text{vor})} \boldsymbol{p}_i = \sum_k^{(\text{nach})} \boldsymbol{p}_k; \qquad \sum_i^{(\text{vor})} E_i = \sum_k^{(\text{nach})} E_k.$$

Dabei sind für die Impulse und Energien die exakten relativistischen Ausdrücke zu benutzen; Besonderheiten nichtrelativistischer Prozesse werden noch diskutiert (↑ 2.1.3.7.). In den Energien sind die Ruhenergien der Teilchen enthalten. Da die Massen der Teilchen vor und nach dem „Stoß" nun nicht mehr übereinzustimmen brauchen, können die Ruhenergien mc^2 aus der allgemeinen Energiebilanz nicht herausgenommen werden; sie sind also nicht mehr belanglose Konstanten. Das ist eine sehr wichtige Aussage: *Bei allgemeinen Wechselwirkungsprozessen sind die Ruhenergien der Teilchen an der Energiebilanz beteiligt.*

Ein Teilchen mit der Energie $mc^2/\sqrt{1 - v^2/c^2}$ kann danach als Träger zweier „Energieformen" aufgefaßt werden: der von der Bewegung des Teilchens unabhängigen *Ruhenergie* mc^2, gewissermaßen seiner *inneren Energie*, und der zusätzlichen *kinetischen Energie* als Folge der Bewegung. Ein *Gedankenexperiment* soll die physikalische Bedeutung der Ruhenergie verdeutlichen: Zwei Teilchen mit entgegengesetzt gleich großen Impulsen, also $\boldsymbol{p}_1 + \boldsymbol{p}_2 = \boldsymbol{0}$, werden aufeinander geschossen (Abb. 2.1.3-4).

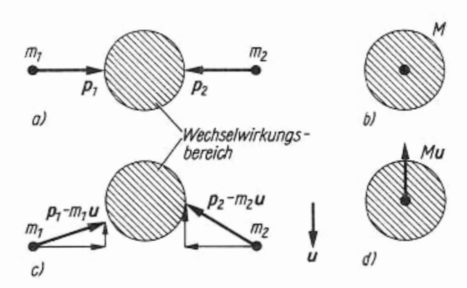


Abb. 2.1.3-4 Verschmelzungsexperiment zur Bestimmung der Ruhmasse; a) und c) vor der Wechselwirkung, b) und d) nach der Verschmelzung, c) und d) im bewegten Bezugssystem; $\boldsymbol{u}$ Relativgeschwindigkeit

Der Mechanismus der Wechselwirkung sei derart, daß zumindest für eine gewisse Zeit beide Teilchen zu einem schwereren „Verbundteilchen" verschmelzen. Dieses hat aufgrund des Impulssatzes keinen Impuls, es ruht also. Aus dem Energiesatz

$$Mc^2 = \frac{m_1 c^2}{\sqrt{1 - v_1^2/c^2}} + \frac{m_2 c^2}{\sqrt{1 - v_2^2/c^2}} = E_1 + E_2$$

wird deutlich, daß die Ruhenergie des neu gebildeten Teilchens außer den Ruhenergien m_1c^2 und m_2c^2 der Ausgangsteilchen auch deren kinetische Energien enthält. Damit ist gezeigt, daß sich kinetische Energie (hier vollständig) in Ruhenergie verwandeln kann. Es ist hier offenbar $M > m_1 + m_2$. Die *Ruhmassen verhalten sich nicht additiv.*

Die *Umwandlungsprozesse zwischen Elementarteilchen* haben eindrucksvoll bestätigt, daß die Ruhenergie von Teilchen eine Energieform ist, die ohne Einschränkungen in die Energiebilanz solcher Prozesse eingeht. Es kann sowohl Ruhenergie völlig neu entstehen, beispielsweise bei der Erzeugung von Teilchen-Antiteilchen-Paaren, als auch völlig in andere Energieformen übergehen, z. B. bei der Zerstrahlung solcher Paare. Bei hochenergetischen Prozessen können die kinetischen Energien groß gegen die Ruhenergien sein.

Der Zusammenhang der inneren Energie E_i eines Teilchens bzw. Körpers mit seiner Ruh-

masse, der durch die berühmte *Einsteinsche Formel*

(8) $E_i = mc^2$

quantitativ formuliert wird, gilt völlig allgemein. Eine komprimierte oder gedehnte Feder hat eine größere Masse als die gleiche Feder im entspannten Zustand. Die Masse eines Körpers vergrößert sich, wenn dieser erhitzt wird. Befindet sich in einem Hohlraum ein elektromagnetisches Strahlenfeld, so trägt die Strahlungsenergie zur Gesamtmasse des Systems bei. Dieser Effekt ist jedoch üblicherweise nicht meßbar, da die Masse des Strahlungsfelds sehr klein gegen die Masse der Wände des Hohlraums ist. Entsprechendes gilt für die Energie der chemischen Bindung, für elastische Energien u. a. Erst in der Kernphysik treten so große Energieänderungen auf, daß die zugehörigen Massen leicht meßbar werden. Die Ruhmasse eines Atomkerns ist kleiner als die Summe der Ruhmassen seiner Bestandteile, der Protonen und Neutronen. Die fehlende Masse ist die Masse der bei Bildung der zusammengesetzten Kerne abgestrahlten Energie.

2.1.3.7. Inelastische Prozesse im nichtrelativistischen Bereich. Häufig sind die Geschwindigkeiten aller Stoßpartner klein gegen die Lichtgeschwindigkeit. Der Impulssatz gilt dann genähert in seiner nichtrelativistischen Form, d. h. mit mv als Impuls eines jeden in die Bilanz eingehenden Teilchens. Gleichzeitig ist die Ruhenergie mc^2 jedes Teilchens groß gegen seine kinetische Energie, die sich genähert als $p^2/2m$ darstellen läßt. In erster Näherung gehen nur die Ruhenergien in die Energiebilanz ein, die nach Weglassen des gemeinsamen Faktors c^2 einfach

$$\sum_i^{(\text{vor})} m_i \approx \sum_k^{(\text{nach})} m_k$$

liefert. Die Summe aller Ruhmassen bleibt bei der Wechselwirkung (genähert) konstant. Dies ist der *Massenerhaltungssatz* der nichtrelativistischen Physik, der häufig als fast selbstverständlich angesehen und daher nicht besonders hervorgehoben bzw. formuliert wird.

Die (Ruh-) Massenerhaltung gilt jedoch exakt nur dann, wenn sich die Summe aller inneren Energien bei der Wechselwirkung nicht ändert. Das ist nur bei der Wechselwirkung von Teilchen ohne innere Struktur und ohne Umwandlungsmöglichkeiten mit Sicherheit erfüllt. Im allgemeinen ändern sich die inneren Energien, und zwar sowohl beim Einzelteilchen als auch in der Summe über alle Teilchen. Der Energiesatz verknüpft die Änderungen der kinetischen Energien mit der Änderung der inneren Energien entsprechend

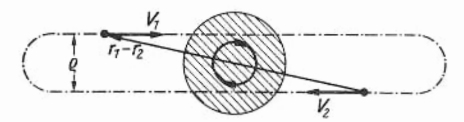


Abb. 2.1.3-5 Zur Bildung des Drehimpulsbegriffs

$$\sum_i^{(\text{vor})} \frac{p_i^2}{2m_i} + \Delta E = \sum_k^{(\text{nach})} \frac{p_k^2}{2m_k}$$

$$\text{mit} \quad \Delta E = \left[\sum_i^{(\text{vor})} m_i - \sum_k^{(\text{nach})} m_k\right] c^2.$$

ΔE ist die beim Prozeß aus dem „Reservoir" der inneren Energien „freiwerdende" Energie, die zur Vergrößerung der kinetischen Energien zur Verfügung steht. $\Delta E < 0$ bedeutet, daß die Summe der inneren Energien anwächst, also kinetische Energie „verbraucht" wird. Häufig sind die Ruhmassen nur so genau bekannt, daß zwar die oben formulierte Massenerhaltung geprüft, aber ΔE praktisch nicht aus den Ruhmassen berechnet werden kann. Dadurch wird ΔE zu einer von den Massen (und damit auch von der Impulsbilanz) scheinbar unabhängigen Größe, die für jeden betrachteten Prozeß zusätzlich anzugeben bzw. durch die Analyse des Stoßvorgangs experimentell zu bestimmen ist.

Ein Prozeß mit genäherter (Ruh-) Massenerhaltung der Einzelteilchen und $\Delta E \neq 0$ wird im engeren Sinne als unelastisch oder inelastisch bezeichnet (im weiteren Sinne sind alle Prozesse mit Ruhmassenänderungen unelastisch.) *Inelastische Stöße* sind eine wichtige Quelle von Informationen über die inneren Energien der Stoßpartner (s. a. 1.1.4.4.).

Die Konzeption, wonach ein Teilchen verschiedene „Anregungszustände" einnehmen kann, zu denen unterschiedliche Werte der inneren Energie gehören, ist an die Voraussetzung gebunden, daß die auftretenden Energiedifferenzen klein gegen die (mittlere) Ruhenergie mc^2 des Teilchens bleiben. Denn zu den unterschiedlichen Energien gehören unterschiedliche Ruhmassen, und „normalerweise" würde man Objekte mit verschiedenen Ruhmassen als verschiedene Teilchen bezeichnen. Offenbar ist die Grenze, von der ab man Teilchen als verschieden bezeichnet, zumindest solange fließend, wie detaillierte Kenntnisse über die innere Struktur der Teilchen fehlen.

Zur Vermeidung von Mißverständnissen sei abschließend betont, daß es bei Benutzung des vollständigen Energieausdrucks $mc^2/\sqrt{1 - v^2/c^2}$ inelastische Prozesse im Sinne einer Verletzung der Energieerhaltung gar nicht gibt. Das gilt auch für makroskopische Prozesse, etwa für den plastischen Stoß zweier Bleikugeln.

2.1.3.8. Drehimpuls. Die beim elastischen Stoß zweier gleicher Teilchen mit entgegengesetzten Geschwindigkeiten $\boldsymbol{V}_1 + \boldsymbol{V}_2 = \boldsymbol{0}$ vorhandenen

Symmetrien sind mit den beiden zum Impuls- und Energiesatz führenden Relationen $V_1' + V_2' = 0$ und $V_1' = V_1$ (↑ 2.1.3.2.) nicht erschöpft. Die zueinander parallelen Geraden, auf denen sich beide Teilchen vor dem Stoß mit den Geschwindigkeiten V_1 und V_2 bewegen, definieren (außer beim zentralen Stoß) eine Ebene, einen Drehsinn in dieser Ebene und einen Stoßabstand ϱ, der häufig als *Stoßparameter* bezeichnet wird (Abb. 2.1.3-5). Analog ist die Situation nach dem Stoß. Falls die Teilchen selbst nicht eine Richtung bzw. eine Achse mit sich tragen und damit besonders auszeichnen (↑ 2.1.3.9.), kann die Orientierung der „Stoßebene" wegen der Gleichwertigkeit aller Raumrichtungen beim Stoß nicht verändert werden. Der Drehsinn bleibt ebenfalls unverändert. Auch der Stoßparameter ϱ ist nach dem Stoß derselbe wie vorher, also $\varrho = \varrho'$. Die auf dem Vergleich von Prozeß und Umkehrprozeß beruhende Argumentation ist analog zu derjenigen, die für $V_1 = V_1'$ gegeben wurde (↑ 2.1.3.2.).

Die drei neuen Aussagen über die *Invarianz von Stoßebene, -drehsinn und -parameter* können unter Beachtung von $V_1 = V_1'$ in der Vektorgleichung $(r_1 - r_2) \times V_1 = (r_1' - r_2') \times V_1'$ zusammengefaßt werden. Die Bildung
$(r_1 - r_2) \times V_1$ selbst kann
wegen $V_2 = -V_1$ auch in der Form
$r_1 \times V_1 + r_2 \times V_2$ geschrieben werden, wobei nun die Orte und Geschwindigkeiten beider Stoßpartner in völlig symmetrischer Weise auftreten.

Allgemein gilt für einen beliebigen Stoß zwischen spinlosen Teilchen (↑ 2.1.3.9.) mit den Massen m_1 und m_2 der sogenannte *Drehimpulssatz* $r_1 \times p_1 + r_2 \times p_2 = r_1' \times p_1' + r_2' \times p_2'$. Der Vektor $L = r \times p$ heißt Drehimpuls, genauer *Bahndrehimpuls* eines Teilchens. Er hängt – im Gegensatz zum Impuls oder zur Energie – von der Wahl des Bezugspunkts ab.

2.1.3.9. Spin. Impuls- und Energiesatz gelten völlig allgemein für beliebige Stoßprozesse (↑ 2.1.3.6.). Der bisher formulierte Drehimpulssatz hat diesen Grad von Allgemeinheit noch nicht. Betrachtet man etwa nochmals die Bildung eines Verbundteilchens beim Stoß (↑ Abb. 2.1.3-4), so besitzt dieses als Ganzes ruhende Teilchen gar keinen Bahndrehimpuls. Es kann aber eine *innere Rotation* zeigen, wenn die beiden „eingefangenen" Teilchen „umeinanderkreisen". Mit der inneren Rotation oder Eigenrotation eines Teilchens ist ein *Eigendrehimpuls* verbunden, der von der Geschwindigkeit des Teilchens völlig unabhängig ist. Dieser Eigendrehimpuls wird *Spin* genannt [*to spin* engl., sich drehen, rotieren]. In dem diskutierten Verschmelzungsexperiment ist der Spin s des gebildeten Teilchens gleich der Vektorsumme der Bahndrehimpulse der Ausgangsteilchen, falls diese selbst keinen Spin besitzen. Es gilt also $s = L_1 + L_2$. Bewegt sich das gebildete Verbundteilchen, so ist es sowohl Träger von Spin- als auch von Bahndrehimpuls, sein *Gesamtdrehimpuls* ist dann $L_{ges} = r \times p + s$. Zerfällt es, so muß die Vektorsumme der Bahn- und Spindrehimpulse der gebildeten Teilchen mit L_{ges} übereinstimmen.

Der Spin als innerer Drehimpuls ist eine Größe, die völlig analog zur inneren Energie eines Teilchens ist. (Es wäre korrekt, wenn auch nicht üblich, den Spin als Ruhdrehimpuls analog zur Ruhenergie zu bezeichnen.) Falls es möglich ist, das betrachtete Teilchen als aus einfacheren Teilchen zusammengesetzt aufzufassen, kann sein Spin als Summe aller Bahn- und Spindrehimpulse seiner Bestandteile erklärt bzw. berechnet werden. Dabei sind die Bahndrehimpulse in demjenigen Bezugssystem zu nehmen, in dem das Gesamtsystem als Ganzes, genauer dessen Schwerpunkt, ruht. Die skizzierte Zurückführung ist jedoch weder immer zweckmäßig noch überhaupt immer möglich. So besitzen auch Elementarteilchen, die sicher nicht in einfachem Sinn als „zusammengesetzt" angesehen werden können, häufig einen für sie charakteristischen Spin.

Allgemein gilt der Drehimpulserhaltungssatz nur bei Einbeziehung der Spindrehimpulse aller Stoßpartner. Dennoch ist die Kopplung zwischen Spin- und Bahndrehimpulsen oft sehr schwach, so daß (genähert) je ein Erhaltungssatz für die Spin- und Bahndrehimpulse gilt bzw. die Einzelspins überhaupt ungeändert bleiben. Jeder Effekt, der eine Kopplung beider Drehimpulsarten bewirkt, wird als *Spin-Bahn-Kopplung* bezeichnet. Beispielsweise führen magnetische Wechselwirkungen zwischen elektrisch geladenen Teilchen oder Reibungskräfte beim Stoß rotierender Billardkugeln zu einer solchen Kopplung. Prozesse, die mit einer Änderung der inneren Energien, also mit Teilchenumwandlungen bzw. unelastischen Prozessen im Sinne von 2.1.3.7. verbunden sind, führen im allgemeinen auch zu Spinänderungen.

2.1.4. Kinematik von Teilchen

2.1.4.1. Geschwindigkeit und Beschleunigung. Im vorigen Abschnitt wurden bei der Untersuchung von Stoßprozessen wichtige Bilanzaussagen formuliert, die nur den Vergleich der Bewegungszustände vor und nach der Wechselwirkung, vor und nach dem Stoß, betreffen. Die Bewegungen der Teilchen sind außerhalb des Wechselwirkungsbereichs sehr einfach, nämlich freie Bewegungen mit konstanter Geschwindigkeit. Im allgemeinen, beispielsweise während des Stoßvorgangs, ist jedoch die Teil-

chenbewegung komplizierter: Die Bahnkurve ist gekrümmt und die Bahngeschwindigkeit ändert sich.

Eine vollständige Beschreibung der Bahnbewegung eines Teilchens erfordert, seinen Ort (bezüglich eines gewählten Bezugssystems) als Funktion der Zeit anzugeben. Das kann in kompakter Weise durch Benutzung des vom Bezugspunkt zum Ort des Teilchens führenden *Ortsvektors* $\boldsymbol{r}$ geschehen (↑ 2.1.2.6.). Durch $\boldsymbol{r} = \boldsymbol{r}(t)$ wird dann die Teilchenbewegung beschrieben. Entsprechend dem allgemeinen Funktionsbegriff ist damit nicht mehr als die paarweise Zuordnung von Raumpunkten und Zeiten, d. h. von Werten der Zeitkoordinate t (↑ 2.1.2.2.) gemeint.

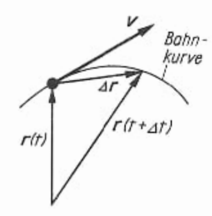


Abb. 2.1.4-1 Zur Definition der Momentangeschwindigkeit

Die *momentane Geschwindigkeit* $\boldsymbol{v}$ ist als Grenzwert des aus der Änderung $\Delta\boldsymbol{r}$ des Ortsvektors und dem zugehörigen Zeitintervall Δt gebildeten Quotienten definiert. Dieser Grenzwert definiert die Zeitableitung $\dot{\boldsymbol{r}} = \mathrm{d}\boldsymbol{r}/\mathrm{d}t$ der Funktion $\boldsymbol{r}(t)$. $\Delta\boldsymbol{r}$ und damit auch $\Delta\boldsymbol{r}/\Delta t$ hat die Richtung der Sekante (Abb. 2.1.4-1) und $\boldsymbol{v}$ nach dem Grenzübergang die *Richtung der Tangente* zur Bahnkurve im Punkt $\boldsymbol{r}(t)$. Der Betrag $|\Delta\boldsymbol{r}|$ hat die Bedeutung der Sekantenlänge, die beim Grenzübergang gleich der infinitesimalen Bogenlänge $\mathrm{d}s$ wird. Daher stellt der Betrag $v = \mathrm{d}s/\mathrm{d}t$ des Geschwindigkeitsvektors die *Bahngeschwindigkeit* dar. Der Vektor $\boldsymbol{v}$ vereint also in sich Bahngeschwindigkeit und Bewegungsrichtung des Teilchens.

Die *momentane Beschleunigung* $\boldsymbol{a}$ ist als Grenzwert des aus der Änderung $\Delta\boldsymbol{v}$ des Geschwindigkeitsvektors und dem zugehörigen Zeitintervall Δt gebildeten Quotienten definiert.

$\boldsymbol{a}$ ist die Zeitableitung $\mathrm{d}\boldsymbol{v}/\mathrm{d}t$ der Funktion $\boldsymbol{v}(t)$. Die Geschwindigkeitsänderung $\Delta\boldsymbol{v}$ ist sowohl dann ungleich Null, wenn sich bei konstanter Bewegungsrichtung nur die Bahngeschwindigkeit ändert, als auch dann, wenn sich bei konstanter Bahngeschwindigkeit nur die Bewegungsrichtung ändert. Im allgemeinen treten beide Spezialfälle kombiniert auf. Im ersten Fall bewegt sich das Teilchen auf einer Geraden, und die Beschleunigung $\boldsymbol{a}$ hat die gleiche (bei wachsendem v) oder die entgegengesetzte Richtung (bei abnehmendem v) wie die Geschwindigkeit $\boldsymbol{v}$. Im zweiten Fall kann es sich noch um eine beliebige gekrümmte Bahnkurve handeln, die von dem Teilchen mit konstanter Bahngeschwindigkeit durchlaufen wird. Die Beschleunigung $\boldsymbol{a}$ steht dann senkrecht auf der Bahn. Eine Beschleunigungskomponente in Bewegungsrichtung bedeutet eine Änderung der Bahngeschwindigkeit, während eine Beschleunigung senkrecht zur Bewegungsrichtung mit einer Bahnkrümmung verbunden ist.

2.1.4.2. Kreisbewegung. Besonders einfach und wichtig ist die Bewegung mit konstanter Bahngeschwindigkeit v auf einer Kreisbahn mit dem Radius R. Der Beschleunigungsvektor ist dann zum Kreismittelpunkt hin gerichtet; sein Betrag ist $a = v^2/R$. Die Größe $\omega \equiv \mathrm{d}\varphi/\mathrm{d}t$ heißt *Winkelgeschwindigkeit*, da sie die zeitliche Änderung des Drehwinkels charakterisiert. Entsprechend $v = R\omega$ und $a = v\omega$ können Bahngeschwindigkeit und Beschleunigung unmittelbar durch die Winkelgeschwindigkeit ausgedrückt werden.

2.1.4.3. Harmonische Schwingung. Häufig schwingt ein Teilchen um eine mittlere Lage. Im einfachsten Fall, nämlich bei einer *linearen Schwingung*, bewegt sich das Teilchen dabei auf einer Geraden, die zur mathematischen Beschreibung der Bewegung als x-Achse gewählt werden kann. Die Schwingung ist harmonisch, wenn $x(t)$ der Gleichung $x(t) = A\cos(\omega t - \alpha)$ genügt (Abb. 2.1.4-2). Das Teilchen bewegt sich zwischen den beiden Punkten $x = -A$ und $x = +A$ um den mittleren Punkt $x = 0$. A heißt *Amplitude* der Schwingung. Die Auslenkungen erfolgen von $x = 0$ aus völlig symmetrisch nach beiden Seiten. Für eine volle Schwingung benötigt das Teilchen die als *Schwingungsdauer* T bezeichnete Zeit. In dieser Zeit ändert sich das Argument $(\omega t - \alpha)$ der cos-Funktion um 2π, also $\omega T = 2\pi$ oder $\omega = 2\pi/T = 2\pi\nu$. Diese Relation verdeutlicht die Bedeutung der Größe ω, die als *Kreisfrequenz* bezeichnet wird. $\nu = 1/T$ ist die *Schwingungsfrequenz*. Sie gibt die Anzahl der Schwingungen pro Sekunde an. Bei der Angabe von Frequenzen ist für die Einheit $1\,\mathrm{s}^{-1}$ auch die Bezeichnung 1 Hz (Hertz) üblich. Die dimensionslose Größe α bestimmt lediglich die Phasenlage der Schwingung: Für $\alpha = 0$ befindet sich das Teilchen zur Zeit $t = 0$ am Umkehrpunkt $x = A$, für $\alpha = \pi/2$ entsteht $x(t) = A\sin\omega t$;

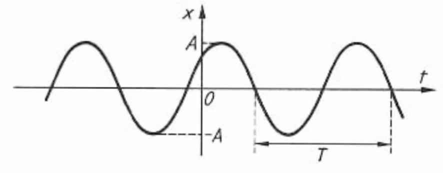


Abb. 2.1.4-2 Darstellung einer harmonischen Schwingung zwischen den Punkten $-A$ und $+A$

das Teilchen durchläuft bei $t = 0$ gerade die mittlere Position $x = 0$. Die Phasenlage einer einzelnen Schwingung ist im allgemeinen belanglos, während bei der Überlagerung mehrerer Schwingungen die Phasendifferenzen wichtig werden.

Die *Überlagerung zweier harmonischer Schwingungen* mit gleicher Frequenz und gleicher Mittellage, aber in unterschiedlichen Richtungen liefert eine *elliptische Schwingung*, bei der das Teilchen in der durch die beiden Schwingungsrichtungen bestimmten Ebene auf einer Ellipse umläuft. Zirkulare und lineare Schwingungen sind Spezialfälle dieses allgemeinen Schwingungstyps. Auch die Überlagerung von drei (oder mehr) linearen Schwingungen liefert eine elliptische Schwingung.

Harmonische Schwingungen treten immer dann auf, wenn auf das Teilchen eine Kraft (↑ 2.1.5.1.) wirkt, die zu dem als Mittellage bezeichneten Punkt O gerichtet ist und deren Stärke F linear mit dem Abstand des Teilchens von O anwächst. Ein solches System heißt *harmonischer Oszillator*. Im Punkt O verschwindet die Kraft. Wenn das Teilchen dort zu einem Zeitpunkt ruht, so bleibt es weiterhin in Ruhe. Der Punkt O kann in diesem Sinne als *Ruhelage* bezeichnet werden.

2.1.5. Dynamik von Teilchen

2.1.5.1. Kraft und Bewegungsgleichung. Beim Stoß zweier Teilchen ändern sich die Geschwindigkeiten, sobald beide Teilchen sich gegenseitig beeinflussen, in Wechselwirkung miteinander treten. Dabei verhalten sich die Geschwindigkeitsänderungen beim elastischen Stoß umgekehrt wie die Teilchenmassen (↑ 2.1.3.3.), während die Impulsänderungen aufgrund des Impulssatzes entgegengesetzt gleich sind. Benutzt man also die Änderung des Impulses als Maß für die Einwirkung auf ein Teilchen, so wirken beim Stoß beide Stoßpartner gleich stark aufeinander. Das ist eine sinnvolle Aussage, die als *Gegenwirkungsprinzip* bezeichnet und oft mit dem Stichwort *actio = reactio* charakterisiert wird.

Das genaue Maß für die gerichtete Einwirkung auf ein Teilchen ist der *Kraftvektor* $\boldsymbol{F}$. Dem Verhalten beim Stoß entsprechend wird die Kraft so definiert, daß sie mit der zeitlichen Ableitung des Teilchenimpulses übereinstimmt, also

$$(9)\quad \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = \boldsymbol{F} \begin{cases} \nearrow\ m\dfrac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = m\boldsymbol{a} = \boldsymbol{F} & \text{für } v \ll c \\ \searrow\ m\dfrac{\mathrm{d}}{\mathrm{d}t}\,\dfrac{\boldsymbol{v}}{1 - v^2/c^2} = \boldsymbol{F} & \text{allgemein;} \end{cases}$$

dabei ist m die (Ruh-) Masse. Aus dem Impulssatz folgt, daß bei der Wechselwirkung zweier Teilchen die auf beide wirkenden Kräfte entgegengesetzt gleich sind (Abb. 2.1.5-1).

In einfachen Experimenten kann man prüfen, ob die durch $\mathrm{d}\boldsymbol{p}/\mathrm{d}t$ gegebene Kraftdefinition physikalisch korrekt ist. Realisiert man beispielsweise eine äußere Einwirkung auf ein makroskopisches Teilchen, indem man die

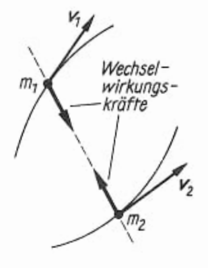


Abb. 2.1.5-1 Zur Bedeutung des Gegenwirkungsprinzips

Kraftübertragung über eine Feder vornimmt, so kann man die Konstanz der Einwirkung, also die Konstanz von $\boldsymbol{F}$, unabhängig von der gegebenen Definition dadurch kontrollieren, daß die Feder stets in gleicher Weise beansprucht, also um das gleiche Stück gedehnt oder zusammengedrückt ist. Das Experiment zeigt dann, daß sich bei verschiedenen Teilchen und gleicher Kraft die Beschleunigungen bei $v \ll c$ wirklich umgekehrt wie die Massen verhalten.

Äußere Einwirkungen, also Kräfte auf ein Teilchen, treten nicht nur bei Stößen auf, sondern beispielsweise auch durch die Wirkung reibender Medien (Reibungskräfte) oder als Folge von elektrischen und magnetischen Feldern, die nicht einem oder mehreren anderen Teilchen in einfacher Weise zuordenbar sein müssen.

Der Zusammenhang [↑ 2.1. (9)] zwischen Impulsänderung und Kraft wird als *Bewegungsgleichung* bezeichnet. $m\boldsymbol{a} = \boldsymbol{F}$ ist also die Bewegungsgleichung eines nichtrelativistischen Teilchens. In der relativistischen Bewegungsgleichung tritt die geschwindigkeitsabhängige Masse (↑ Abb. 2.1.3-3) auf. Da diese Masse für $v \to c$ unbegrenzt anwächst, kann ein Teilchen (mit $m \neq 0$) niemals auf Lichtgeschwindigkeit beschleunigt werden. Die relativistische Geschwindigkeitsabhängigkeit der Masse wird vom Experiment vollständig bestätigt, vor allem durch die schnelle Bewegung geladener Teilchen (Elektronen, Protonen) in Teilchenbeschleunigern (↑ 3.1.).

2.1.5.2. Kraftgesetz. Die auf ein Teilchen einwirkende Kraft kann abhängen:

vom Ort $\boldsymbol{r}$, an dem sich das Teilchen befindet, z. B. in einem inhomogenen elektrischen Feld oder bei der sich mit dem Abstand vom Erdmittelpunkt ändernden Erdanziehung;

von der Geschwindigkeit $\boldsymbol{v}$ des bewegten Teil-

chens, z. B. bei Reibungskräften oder bei der Kraft auf eine bewegte elektrische Ladung im Magnetfeld;
von der Zeit t. Damit ist nicht gemeint, daß auf das Teilchen zu verschiedenen Zeiten einfach deshalb verschiedene Kräfte wirken, weil sich das Teilchen an einem anderen Ort befindet oder sich mit einer anderen Geschwindigkeit bewegt; vielmehr werden dadurch Fälle beschrieben, in denen sich die Kraft auch bei festen Werten von Ort und Geschwindigkeit ändern würde, etwa durch Ein- oder Ausschalten elektrischer Felder, oder dadurch, daß das Teilchen mit einem oder mehreren anderen Teilchen wechselwirkt, die sich in bekannter Weise bewegen.

2.1.5.3. Determinismus. Bei bekanntem Kraftgesetz können nur noch Ort und Geschwindigkeit des Teilchens zu einem einzigen Zeitpunkt frei gewählt werden: Das Teilchen „startet" an einem „Ausgangs"punkt mit seiner „Anfangs"-geschwindigkeit. Diese *Anfangswerte* können jedoch auch einfach diejenigen Werte sein, die sich bei einer ablaufenden Bewegung zur gewählten „Anfangs"zeit einstellen. Der gesamte Bewegungsablauf ist dann eindeutig bestimmt. Diese Determiniertheit erstreckt sich nicht nur auf das zukünftige, sondern auch auf das vergangene Verhalten des mechanischen Systems.

2.1.5.4. Kinetische Energie und Arbeit. Zusammen mit dem Impuls ändert sich im allgemeinen auch die kinetische Energie des Teilchens. Da diese eine eindeutige Funktion des Impulses (bzw. der Geschwindigkeit) ist (↑ 2.1.3.4. und 2.1.3.5.), ist ihre Änderung ebenfalls durch die Bewegungsgleichung bestimmt. Die zeitliche Ableitung der kinetischen Energie ist $\mathrm{d}E_k/\mathrm{d}t = \mathrm{d}E_k/\mathrm{d}p \cdot \mathrm{d}p/\mathrm{d}t = v\,\mathrm{d}p/\mathrm{d}t$. Da $\mathrm{d}E_k/\mathrm{d}p = v$ sowohl im nichtrealtivistischen Fall, also mit $p = mv$ und $E_k = p^2/2m$, als auch im relativistischen Fall (↑ 2.1.3.5.) gilt, ist die obige Ableitung allgemein gültig. Da die im relativistischen Energieausdruck $E = c\sqrt{(mc^2) + p^2}$ enthaltene Ruhenergie mc^2 bei der Bewegung des Teilchens unter Krafteinwirkung konstant bleibt, ist $\mathrm{d}E/\mathrm{d}t = \mathrm{d}E_k/\mathrm{d}t$ wie behauptet wirklich die zeitliche Ableitung der kinetischen Energie.
Mit der Bewegungsgleichung kann $\mathrm{d}E_k/\mathrm{d}t$ in der Form $\mathrm{d}E_k/\mathrm{d}t = v\,\mathrm{d}p/\mathrm{d}t = vF_t = \boldsymbol{v}\boldsymbol{F}$ dargestellt werden. Nur die tangential zur Bahnkurve orientierte Komponente F_t der Kraft $\boldsymbol{F}$ trägt zur Änderung der kinetischen Energie bei.
Da die Geschwindigkeit $\boldsymbol{v} = \mathrm{d}\boldsymbol{r}/\mathrm{d}t$ unmittelbar mit der Verschiebung $\mathrm{d}\boldsymbol{r}$ des Teilchens längs seiner Bahn zusammenhängt, kann das obige Ergebnis auch in der Form $\mathrm{d}E_k = \boldsymbol{F}\,\mathrm{d}\boldsymbol{r}$ dargestellt werden. Das Skalarprodukt $\boldsymbol{F}\,\mathrm{d}\boldsymbol{r}$ ist die von der Kraft $\boldsymbol{F}$ auf dem Wegstück $\mathrm{d}\boldsymbol{r}$ verrichtete *Arbeit*. Die Änderung der kinetischen Energie stimmt also mit der am Teilchen verrichteten Arbeit überein. Die Arbeit pro Zeiteinheit $\boldsymbol{F}\,\mathrm{d}\boldsymbol{r}/\mathrm{d}t = \boldsymbol{F}\boldsymbol{v}$ wird *Leistung* genannt.
Als eindimensionales *Beispiel* wird die von der Federkraft an einem schwingenden Teilchen verrichtete Arbeit betrachtet. Es sei $x = 0$ die Ruhelage, bei der die Feder entspannt ist. Bei hinreichend kleinen Auslenkungen x aus der Ruhelage ist die rücktreibende Kraft $F = -ax$ der Auslenkung proportional, a ist die Federkonstante. Bewegt sich (schwingt) das Teilchen von der Ruhelage $x = 0$ bis zum Punkt x, so verrichtet die Kraft an dem Teilchen die Arbeit
$W = \int_0^x F(x)\,\mathrm{d}x = -a\int_0^x x'\,\mathrm{d}x' = -\frac{a}{2}x^2$. Die Arbeit ist unabhängig vom Vorzeichen von x, d. h. davon, ob die Feder um $|x|$ gedehnt oder zusammengedrückt wird. Sie ist negativ, so daß die kinetische Energie des Teilchens mit wachsender Auslenkung abnimmt. Hat das Teilchen beim Durchgang durch die Ruhelage die Geschwindigkeit v_0, also (nichtrelativistisch) die kinetische Energie $mv_0^2/2$, so kann es nur bis zu der durch $mv_0^2/2 = aA^2/2$ bestimmten maximalen Auslenkung schwingen: A ist die Amplitude der Schwingung (↑ 2.1.4.3.).

2.1.5.5. Potentielle Energie und Gesamtenergie. Die Arbeit bei dem an einer Feder schwingenden Teilchen (↑ 2.1.5.4.) hängt nur von den beiden Endpunkten des betrachteten Bahnstücks ab, nicht aber von dem genauen Bewegungsablauf, also z. B. von der Geschwindigkeit des Teilchens. Ein solches Verhalten der Arbeit ist sehr häufig. Es ist nicht zufällig, sondern hängt eng mit dem allgemeinen Energieprinzip (↑ 2.2.5.1.) zusammen.
Danach muß die Erhöhung bzw. Verringerung der kinetischen Energie eines Teilchens immer mit einer betragsmäßig gleichgroßen Verringerung bzw. Erhöhung der Energie anderer physikalischer Objekte wie weiterer Teilchen, Felder, Federn verbunden sein. Im obigen Beispiel ist das zweite Objekt die Feder, seine Energie deren elastische Energie. Im allgemeinen wird diese Energie allerdings nicht nur von der jeweiligen Lage des Teilchens, sondern darüber hinaus explizit von den inneren Freiheitsgraden des zweiten Systems abhängen. Falls jedoch die für die Bewegung des Teilchens charakteristischen Zeiten, z. B. die Schwingungsdauer der Federschwingung, sehr viel größer sind als die charakteristischen Zeiten des zweiten Systems, z. B. als die Relaxationszeit für elastisches Gleichgewicht bei der Feder, als die Laufzeit elektromagnetischer Wellen bei Kopplung mit elektromagnetischen Feldern, so hängt der gesamte Zustand und damit auch die

Energie E' des zweiten Systems nur von der momentanen Lage des Teilchens ab: $E' = E'(\boldsymbol{r})$. Die *Gesamtenergie* ist die Summe aus der kinetischen Energie $E_k(\boldsymbol{p})$ des Teilchens und der Energie E' des angekoppelten Systems. Der *allgemeine Energiesatz* fordert: $E_g = E_k(\boldsymbol{p}) + E' = \text{const.}$ Falls die Voraussetzung, daß E' wirklich nur vom Ort $\boldsymbol{r}$ des Teilchens abhängt, erfüllt ist, nennt man $E' = E_p(\boldsymbol{r})$ die *potentielle Energie* des Teilchens und die Aussage

$$(10) \quad E_g = E_k(\boldsymbol{p}) + E_p(\boldsymbol{r}) = \text{const.}$$

den *mechanischen Energiesatz*. Es wird deutlich, daß die potentielle Energie keine physikalisch selbständige Energieform darstellt. Da ihre Einführung der obigen Diskussion von E' entsprechend auf hinreichend langsame Bewegungen beschränkt ist, ist der Begriff der potentiellen Energie für die Theorie nichtrelativistischer Teilchenbewegungen typisch.

Da die Änderung von $E_k(\boldsymbol{p})$ gleich der Arbeit ist, muß die Änderung der potentiellen Energie gleich der negativen Arbeit sein, also $\mathrm{d}E_p = -\boldsymbol{F}\,\mathrm{d}\boldsymbol{r}$ und $E_p(\boldsymbol{r}_2) - E_p(\boldsymbol{r}_1) = -\int_{\boldsymbol{r}_1}^{\boldsymbol{r}_2} \boldsymbol{F}\,\mathrm{d}\boldsymbol{r}$. Daraus folgt, daß die Kraft nur vom Ort $\boldsymbol{r}$ abhängen darf, da anderenfalls die Arbeit sicher nicht nur von Ausgangs- und Endpunkt abhängt.

Die bei zwei Teilchen auftretende *potentielle Energie der gegenseitigen Wechselwirkung* hängt nur vom Abstand $|\boldsymbol{r}_1 - \boldsymbol{r}_2|$ beider Teilchen ab. Die Gesamtenergie besteht aus den kinetischen Energien beider Teilchen und der *Wechselwirkungsenergie* E_p^w. Falls auf die Teilchen außerdem äußere, d. h. nicht von jeweils anderen Teilchen herrührende Kräfte wirken, so sind die zugehörigen potentiellen Energien mit in die Gesamtenergie aufzunehmen. Bei mehr als zwei Teilchen tritt für jedes Teilchenpaar eine Wechselwirkungsenergie auf, bei drei Teilchen mit den Indizes 1, 2 und 3 also die für die Paare (1 2), (1 3) und (2 3). Bei N Teilchen gibt es $N(N-1)/2$ verschiedene Paare. Von größter Bedeutung für die Struktur der Materie ist die elektrostatische Wechselwirkung zwischen geladenen Teilchen. Es ist $E_p^w = q_1 q_2/4\pi\varepsilon_0|\boldsymbol{r}_1 - \boldsymbol{r}_2|$ das Wechselwirkungspotential zweier Teilchen mit den Ladungen q_1 und q_2 (s. a. 2.2.3.3.).

2.1.5.6. Kraft und Potential. Im eindimensionalen Fall folgt aus $\mathrm{d}E_p(x) = -F\,\mathrm{d}x$ unmittelbar, daß die Kraft $F(x) = -\mathrm{d}E_p/\mathrm{d}x$ aus der potentiellen Energie durch Bildung des negativen Differentialquotienten berechnet werden kann. Will man diesen Umstand hervorheben, daß nämlich aus der potentiellen Energie die Kraft ermittelt werden kann, so nennt man die Funktion $E_p(x)$ bzw. $E_p(\boldsymbol{r})$ das *Potential der Kraft* und bezeichnet sie mit $V(x)$ bzw. $V(\boldsymbol{r})$. Während man also von der potentiellen Energie meist im Hinblick auf eine konkrete Lage oder Bewegung des Teilchens sprechen wird, ist das Potential im ganzen Raum definiert ohne Rücksicht darauf, welche Punkte bei einer konkreten Bewegung wirklich vom Teilchen durchlaufen werden. Geht man von der Kraft aus, so sind durch die von dieser verrichteten Arbeit immer nur Differenzen der potentiellen Energie bzw. des Potentials bestimmt; das Potential $V(\boldsymbol{r})$ enthält also eine frei wählbare Konstante.

Einige *Besonderheiten* sind *im räumlichen Fall* zu beachten. Zunächst folgt aus der Existenz der potentiellen Energie nicht nur, daß ein rein ortsabhängiges Kraftfeld $\boldsymbol{F}(\boldsymbol{r})$ vorliegt, sondern darüber hinaus, daß die Arbeit zwischen zwei beliebigen Punkten $\boldsymbol{r}_1$ und $\boldsymbol{r}_2$ unabhängig von der diese beiden Punkte verbindenden Raumkurve ist. Kehrt das Teilchen auf einer geschlossenen Bahn zu seinem Ausgangspunkt zurück, so verschwindet die Arbeit für einen solchen geschlossenen Weg. Solche Kräfte werden *konservative Kräfte* genannt. Entsprechend dem allgemeinen Energieprinzip erfüllen die wirklich vorkommenden rein ortsabhängigen Kräfte im allgemeinen auch diese Bedingung.

Den Verlauf des Potentials kann man durch *Potentialflächen* veranschaulichen. Diese sind als Flächen definiert, auf denen das Potential konstante Werte hat. Zu jeder Fläche gehört ein ganz bestimmter konstanter Potentialwert. Zu benachbarten Flächen gehören benachbarte Potentialwerte. Bei Verschiebung eines Teilchens innerhalb einer Potentialfläche wird keine Arbeit geleistet. Daraus folgt, daß der Kraftvektor senkrecht auf den Potentialflächen steht. Verschiebt man ein Teilchen von einer Fläche zu einer benachbarten, so ergibt sich die aufzuwendende Arbeit aus der Differenz ΔV der beiden Potentialwerte. Liegen die beiden Flächen bei gegebenem ΔV dicht beieinander, so wird diese Arbeit auf einem kleinen Wegstück geleistet, die Kraft muß deshalb größer als bei weiter auseinanderliegenden Flächen sein.

Da das Potentialfeld im Gegensatz zum vektoriellen Kraftfeld ein skalares Feld und damit einfacher überschaubar und darstellbar ist, arbeitet man bei Existenz eines Potentials fast ausschließlich mit diesem.

2.1.5.7. Bewegung im Potential. Der Ablauf eindimensionaler Bewegungen wird bei gegebenem Potentialverlauf bis auf den frei wählbaren Zeitpunkt, zu dem irgendein bestimmter Bahnpunkt durchlaufen wird, eindeutig durch den Wert der Gesamtenergie bestimmt. Den qualitativen Verlauf der Bewegung $x(t)$ im Potential $V(x)$ übersieht man wie folgt: Da die kinetische Energie $mv^2/2 = E_g - V(x)$ immer positiv sein muß, sind dem Teilchen nur solche Bereiche der x-Achse zugänglich, für die das Potential $V(x)$ kleiner als der vorgegebene

Energiewert E_g ist. Ist das zwischen den Punkten x_1 und x_2 der Fall, in denen die Geschwindigkeit v Null wird, so sind x_1 und x_2 die beiden *Umkehrpunkte* der auf das Gebiet $x_1 \dots x_2$ beschränkten Bewegung (Abb. 2.1.5-2). Zu jedem Punkt x kann aus der Gesamtenergie E_g die Geschwindigkeit $v = \pm\sqrt{2[E_g - V(x)]/m}$ berechnet werden. Die beiden Vorzeichen entsprechen den möglichen Bewegungsrichtungen.

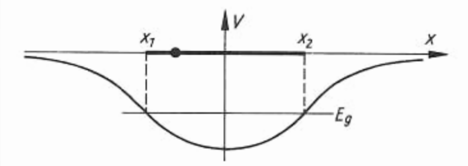


Abb. 2.1.5-2 Schwingung im Potential zwischen zwei Umkehrpunkten

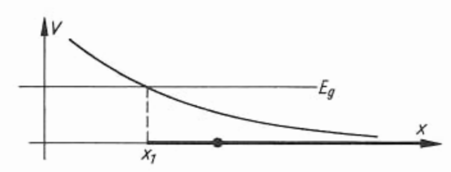


Abb. 2.1.5-3 Reflexion an einer Potentialwand

Insgesamt entsteht ein sich periodisch wiederholender Schwingungsvorgang, bei dem das Teilchen zwischen den beiden Umkehrpunkten hin und her pendelt.

Eine andere typische Bewegung (Abb. 2.1.5-3) hat nur einen Umkehrpunkt x_1. Das Teilchen kommt etwa von $x = +\infty$, läuft bis $x = x_1$ „den Potentialberg hinauf", wird dabei abgebremst, kehrt um und läuft zurück.

In beiden Beispielen ist das Potential im Unendlichen Null gesetzt worden. Falls das überhaupt möglich ist, wird über die freie Konstante im Potential stets so verfügt.

2.1.5.8. Drehmoment. Außer durch seinen Impuls und seine Energie wird ein bewegtes Teilchen durch seinen Drehimpuls (↑ 2.1.3.8., 2.1.3.9.) charakteristiert. Dieser besteht im allgemeinen aus zwei Anteilen: dem Bahndrehimpuls $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ und dem Eigendrehimpuls oder Spin. Der Bahndrehimpuls hängt unmittelbar mit den die Bewegung des Teilchens charakterisierenden Größen $\boldsymbol{r}$ und $\boldsymbol{p}$ bzw. $\boldsymbol{v}$ zusammen. Der erste Anteil seiner zeitlichen Ableitung $\mathrm{d}\boldsymbol{L}/\mathrm{d}t = \dot{\boldsymbol{L}} = \dot{\boldsymbol{r}} \times \boldsymbol{p} + \boldsymbol{r} \times \dot{\boldsymbol{p}}$ verschwindet, da $\dot{\boldsymbol{r}} = \boldsymbol{v}$ und $\boldsymbol{p}$ parallele Vektoren sind. Der zweite Anteil enthält $\dot{\boldsymbol{p}} = \mathrm{d}\boldsymbol{p}/\mathrm{d}t$, das entsprechend der Bewegungsgleichung [↑ 2.1. (9)] gleich der wirkenden Kraft $\boldsymbol{F}$ ist. Das Vektorprodukt $\boldsymbol{M} \equiv \boldsymbol{r} \times \boldsymbol{F}$ aus Orts- und Kraftvektor heißt *Drehmoment*, es bestimmt entsprechend $\dot{\boldsymbol{L}} = \boldsymbol{M}$ die zeitliche Änderung des Drehimpulses.

Diese Aussage ist bei der Untersuchung der Bahnbewegung eines Teilchens eigentlich nur dann interessant, wenn das Drehmoment trotz wirkender Kraft verschwindet, der Drehimpulsvektor also konstant ist. Das Drehmoment verschwindet, wenn diese Kraft auf ein *Kraftzentrum* hin (bei Anziehung) oder von diesem weg (bei Abstoßung) gerichtet ist, weil dann der vom Kraftzentrum ausgehende Ortsvektor parallel oder antiparallel zu $\boldsymbol{F}$ ist. Solche Kräfte heißen *Zentralkräfte*. Die Stärke (der Betrag) der tatsächlich vorkommenden Zentralkräfte hängt nur vom Abstand r ab, es handelt sich um kugelsymmetrische konservative Kraftfelder. Als Gravitationsanziehung beherrschen Kräfte dieses Typs die Bewegung der Planeten um die Sonne, als elektrostatische Kräfte die Bewegung der Elektronen im Atom um den positiv geladenen Kern. Aus der Konstanz des Drehimpulsvektors $\boldsymbol{L}$ folgt:

Die Bewegung erfolgt in derjenigen Ebene, die das Kraftzentrum enthält und auf der $\boldsymbol{L}$ senkrecht steht.

Es gilt der *Flächensatz*: Der Ortsvektor des Teilchens überstreicht in gleichen Zeiten gleiche Flächen (Abb. 2.1.5-4).

Die Kraft $\boldsymbol{F}$ ändert nicht die innere Energie und den inneren Drehimpuls, den Spin $\boldsymbol{s}$. Dieser kann sich unter dem Einfluß eines von der Kraft $\boldsymbol{F}$ völlig unabhängigen Drehmoments $\boldsymbol{M}_s$ ändern entsprechend $\mathrm{d}\boldsymbol{s}/\mathrm{d}t = \boldsymbol{M}_s$. Ein solches Drehmoment tritt z. B. immer auf, wenn sich ein geladenes Teilchen mit *Spin in einem Magnetfeld* befindet (↑ 2.2.3.5.). Diese Drehmomente sind von der Form $\boldsymbol{M}_s = \boldsymbol{\omega}_L \times \boldsymbol{s}$ mit einem zum Magnetfeld $\boldsymbol{B}$ proportionalen Vektor $\boldsymbol{\omega}_L$. Die Vektordifferentialgleichung $\mathrm{d}\boldsymbol{s}/\mathrm{d}t = \boldsymbol{\omega}_L \times \boldsymbol{s}$ bedeutet, daß sich der Spinvektor $\boldsymbol{s}$ mit der Winkelgeschwindigkeit ω_L um die durch $\boldsymbol{\omega}_L$ bestimmte Achse dreht. Diese Erscheinung heißt *Larmorpräzession* (s. a. 2.2.3.5.).

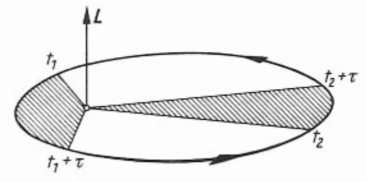


Abb. 2.1.5-4 Drehimpuls und Flächensatz für die Bewegung im Zentralfeld

2.1.6. Einige wichtige Bewegungsprobleme

2.1.6.1. Gedämpfte Schwingungen. Zu den typischsten, alle Gebiete der Physik durchziehenden Erscheinungen gehören gedämpfte und

erzwungene Schwingungen. Diese Erscheinungen können vollständig am mechanischen Modell der eindimensionalen Schwingungen eines Teilchens um eine Ruhelage verstanden werden. Es sei $x(t)$ die Auslenkung aus der Ruhelage. Das Verhalten des reibungsfreien Systems ist durch seine *Eigenfrequenz* ω_0 charakterisiert, die Bewegung ist eine *harmonische Schwingung* (↑ 2.1.4.3.). Durch Hinzunahme einer geschwindigkeitsproportionalen *Reibungskraft* $-\beta v = -\beta\, dx/dt$, z. B. Luftreibung bei einem makroskopischen Teilchen, erhält die Bewegungsgleichung mit der Abkürzung $\gamma \equiv \beta/m$ die für einen gedämpften harmonischen Oszillator typische Gestalt: $d^2x/dt^2 + \gamma\, dx/dt + \omega_0^2 x = 0$.

Falls die Reibungskonstante nicht zu groß ist, genauer für $\gamma/2 < \omega_0$, führt das Teilchen eine *gedämpfte Schwingung* aus, deren Amplitude A sich nach dem Gesetz $A \sim \exp(-\gamma t/2)$ verringert (Abb. 2.1.6-1). Für $\gamma t \gg 1$ bzw. $t \gg \gamma^{-1}$ hat der Schwingungsvorgang praktisch aufgehört. $\tau = \gamma^{-1}$ charakterisiert daher die Zeitdauer, in der das System nach einem einmaligen „Anstoßen" merklich schwingt. Am wichtigsten sind *schwach gedämpfte Systeme*, d. h. $\gamma \ll \omega_0$ bzw. $\tau \gg T$ mit T als Schwingungsdauer, bei denen die Schwingungsamplitude erst im Verlauf sehr vieler Schwingungsperioden merklich abnimmt.

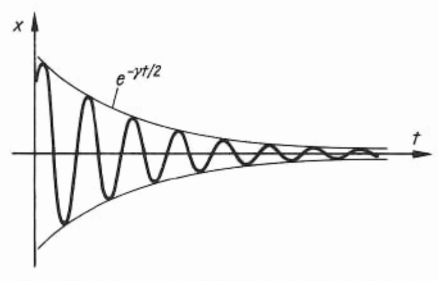


Abb. 2.1.6-1 Graphische Darstellung einer gedämpften Schwingung

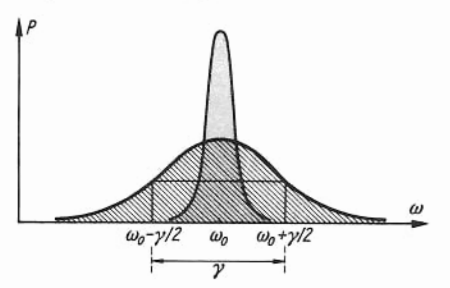


Abb. 2.1.6-2 Frequenzverteilung einer gedämpften Schwingung für zwei verschiedene Dämpfungswerte

Eine gedämpfte Schwingung ist nicht mehr streng harmonisch. Das wird besonders deutlich, wenn man sie nach dem mathematischen Verfahren der *Fourieranalyse* als Überlagerung von rein harmonischen $\sin \omega t$- bzw. $\cos \omega t$-Funktionen darstellt (s. a. 2.3.1.6.). Dabei sind solche Funktionen in der Form eines Integrals über alle Frequenzen zu überlagern, wobei verschiedene Frequenzintervalle unterschiedlich stark beitragen. Das Absolutquadrat des entsprechenden „Gewichtsfaktors" der Fouriertransformierten von $x(t)$ zeigt eine außerordentlich charakteristische Frequenzabhängigkeit (Abb. 2.1.6-2), die formelmäßig durch den Verlauf $P(\omega) \sim [(\omega - \omega_0)^2 + (\gamma/2)^2]^{-1}$ für die *Frequenzverteilung* dargestellt wird. Diese Verteilung hat ihr Maximum bei $\omega = \omega_0$ und ist bei $\omega = \omega_0 \pm \gamma/2$ jeweils auf die Hälfte des Maximalwerts abgesunken. Die Dämpfungskonstante γ bestimmt also die der Frequenzverteilung zugeordnete Halbwertsbreite. Je kleiner γ ist, um so kleiner ist der Bereich von Frequenzen, die wesentlich am Aufbau der Schwingung beteiligt sind, um so „genauer" schwingt das Teilchen mit seiner Eigenfrequenz ω_0. Die Frequenzverteilung einer schwach gedämpften Schwingung zeigt ein ausgeprägtes *Resonanzverhalten*. Die durch die obige Gleichung bestimmte Linienform wird *Lorentzprofil* genannt.

2.1.6.2. Resonanz. Wird das schwingungsfähige System zum Schwingen gezwungen durch eine äußere *erregende Kraft*, die einen rein harmonischen Zeitablauf $F(t) = F_0 \cos \omega t$ mit der (Kreis-) Frequenz ω zeigt, so stellt sich nach einiger Zeit eine Schwingung mit der gleichen, im allgemeinen von der Eigenfrequenz ω_0 verschiedenen Frequenz ω ein. Die dauernd wirkende Kraft führt zu einer *erzwungenen Schwingung* $x(t) = A \cos(\omega t - \varphi)$, die allerdings gegenüber der erregenden Kraft eine frequenzabhängige *Phasenverschiebung* $\varphi(\omega)$ besitzt. Die ebenfalls frequenzabhängige Schwingungsamplitude $A(\omega)$ zeigt *Resonanzverhalten*: Das Amplitudenquadrat $A^2(\omega) \sim P(\omega)$ ist ein „Abbild" der oben diskutierten Frequenzverteilung. Das System reagiert also dann stark auf eine äußere Kraft, wenn deren Frequenz im Bereich seines „eigenen", gewissermaßen „natürlichen" Frequenzbereichs liegt; es reagiert fast gar nicht auf Einwirkungen (Störungen) mit völlig anderer Frequenz. Aus der Reaktion des Systems auf äußere Störungen kann man also auf sein Frequenzspektrum schließen; dies gilt sehr allgemein, nicht nur für eine gedämpfte Schwingung.

Bei Resonanz, d. h. für $\omega = \omega_0$, wird $A^2(\omega_0) \sim P(\omega_0) \sim \gamma^{-2}$ bei einem schwach gedämpften System sehr groß. Die Gesamtenergie des Systems, beim schwingenden Teilchen also die Summe aus kinetischer Energie $mv^2/2$ und

potentieller Energie $m\omega_0^2 x^2/2$, ist proportional zum Amplitudenquadrat und zeigt daher dieselbe Frequenzabhängigkeit.
Durch die Reibung (Dämpfung) wird ständig Schwingungsenergie „verbraucht", also z. B. in Wärme überführt. Da der Energieinhalt der zeitlich unveränderlichen Schwingungsamplitude entsprechend konstant ist, müssen diese Verluste ständig durch die äußere Anregung ersetzt werden: Das zu einer ungedämpften erzwungenen Schwingung angeregte System absorbiert Energie. Die *Verlustleistung* ist gleich der Arbeit, die pro Zeiteinheit von der Reibungskraft verrichtet wird, also gleich $|F_R v| = |\beta v^2| = m\gamma v^2 \sim \gamma A^2$ (↑ 2.1.5.4.). Neben anderen, hier uninteressanten Faktoren ist sie also durch $\gamma P(\omega)$ bestimmt. Das Integral von $\gamma P(\omega)$ über alle Frequenzen ist unabhängig von γ. Regt man daher das System mit einer zeitlich nichtperiodischen Kraft an, die alle Frequenzen innerhalb des Resonanzintervalls „gleichmäßig" enthält, so ist die insgesamt absorbierte Energie unabhängig von der Stärke der Dämpfung.

2.1.6.3. Bewegung im Zentralfeld. Die Bewegung in einem Zentralfeld (↑ 2.1.5.8.) verläuft wie bereits erörtert in einer Ebene. Die Lage und Bewegung des Teilchens kann daher durch ebene *Polarkoordinaten* (Abb. 2.1.6-3) beschrieben werden. Der Drehimpulsbetrag $L = mrv_\perp$ hängt nur von der zum Radius senkrechten Geschwindigkeitskomponente ab, wegen $v_\perp = r\,\mathrm{d}\varphi/\mathrm{d}t$ besteht also der einfache Zusammenhang $L = mr^2\,\mathrm{d}\varphi/\mathrm{d}t$ zwischen Drehimpuls und Winkelgeschwindigkeit $\mathrm{d}\varphi/\mathrm{d}t$. Nur die Änderung des Winkels, also die Drehung des Teilchens um das Zentrum, trägt zu L bei. Dieser Umstand hat zur Bezeichnung „Drehimpuls" geführt.
Das Geschwindigkeitsquadrat $v^2 = (\mathrm{d}r/\mathrm{d}t)^2 + (r\,\mathrm{d}\varphi/\mathrm{d}t)^2$ besteht aus dem Quadrat der radialen Geschwindigkeitskomponente $\mathrm{d}r/\mathrm{d}t$ (Änderung des Abstands r) und $v_\perp^2$. Der von $\mathrm{d}\varphi/\mathrm{d}t$ herrührende Winkelanteil kann durch den konstanten Drehimpuls ausgedrückt werden. Die *Gesamtenergie* erhält auf diese Weise die Form $E_g = \frac{m}{2}v^2 + V(r) = \frac{m}{2}\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 + \frac{L^2}{2mr^2} + V(r)$, der Drehwinkel ist also eliminiert. $V(r)$ ist das Potential (↑ 2.1.5.6.) des Zentralfelds.
Dieser Energieausdruck hat die gleiche Form wie bei einer eindimensionalen Bewegung mit der Koordinate $r(t)$, die im *Effektivpotential* $V_{\mathrm{eff}}^{(r)} \equiv V(r) + \frac{L^2}{2mr^2} = V(r) + V_Z(r)$ abläuft.
Der Term $L^2/2mr^2$ ist stets positiv und beschreibt ein mit dem Quadrat des Drehimpulses anwachsendes Abstoßungspotential. Er wird als *Zentrifugalpotential* V_Z bezeichnet. Tatsächlich stimmt die diesem Potential zugeordnete Kraft $F_Z = -\mathrm{d}V_Z/\mathrm{d}r$ mit der Zentrifugalkraft $m\omega^2 r$ überein, wenn man $\omega = \mathrm{d}\varphi/\mathrm{d}t$ beachtet und $\mathrm{d}\varphi/\mathrm{d}t$ durch L ausdrückt. Auch physikalisch ist die Interpretation korrekt, denn die durch formale Rechnung erreichte Elimination der Winkeländerung kann physikalisch gerade durch Benutzung eines mit der gleichen Winkelgeschwindigkeit wie das Teilchen rotierenden Bezugssystems realisiert werden. Ein Beobachter in diesem System registriert nur die radiale Bewegung, also die Änderung des Abstands des Teilchens vom Kraftzentrum, und gleichzeitig tritt für ihn als rotierenden Beobachter die Zentrifugalkraft auf.
Das Effektivpotential wird für kleine Abstände r stets (außer bei $L = 0$, das ist bei einer rein radialen, also linearen Bewegung) von dem für $r \to 0$ stark anwachsenden und bei $r = 0$ singulären Zentrifugalpotential bestimmt. Mit

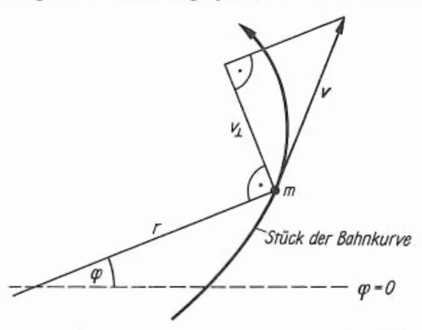


Abb. 2.1.6-3 Polarkoordinaten und Zerlegung der Geschwindigkeit bei einer ebenen Bewegung

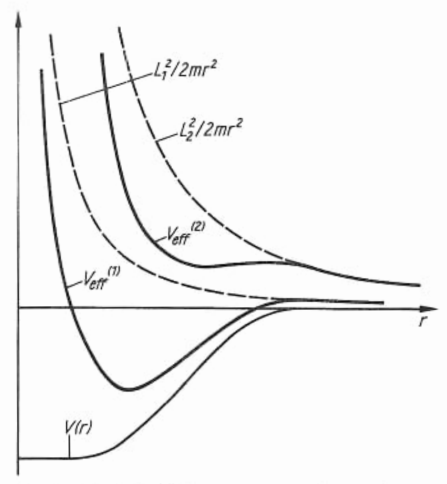


Abb. 2.1.6-4 Effektivpotential für zwei verschiedene Drehimpulse; $L_2 > L_1$

wachsendem Drehimpuls L kann sich auch der qualitative Verlauf des Effektivpotentials verändern (Abb. 2.1.6-4). So ist bei dem dargestellten Beispiel das eigentliche Potential V das eines anziehenden Kraftfelds. Für den kleineren Drehimpuls L_1 zeigt das Effektivpotential zentrifugale Abstoßung bei kleinen und von $V(r)$ herrührende Anziehung bei größeren Abständen. Die sich ausbildende Potentialmulde erlaubt die Ausbildung gebundener Zustände (s. u.), bei denen sich das Teilchen nur bis zu einem maximalen Abstand vom Kraftzentrum entfernen kann. Dagegen ist V_{eff} beim größeren Drehimpuls L_2 ein reines Abstoßungspotential, und gebundene Bewegungen sind unmöglich. Das Potential $V(r)$ bzw. die zugehörige Kraft ist dann also zu schwach, um entgegen den bei großen Drehimpulsen auftretenden starken Zentrifugalkräften eine Bindung an das Zentrum aufrechtzuerhalten.

Die radiale Bewegung kann genauso wie bei einem eindimensionalen Problem (↑ 2.1.5.7.) diskutiert werden. Beim Vergleich von Bewegungen mit gleichem Drehimpuls, also auch gleichem Effektivpotential, ist der Wert der

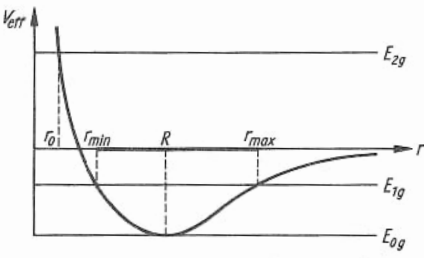


Abb. 2.1.6-5 Bewegungen mit verschiedener Energie bei festem Drehimpuls

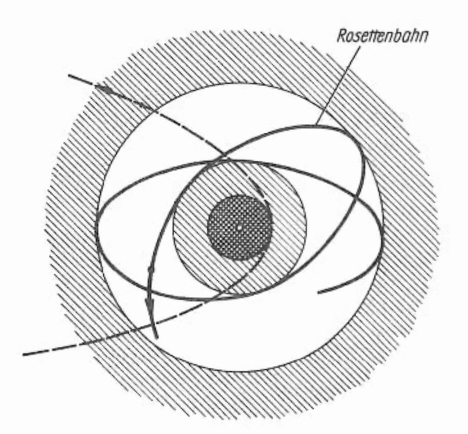


Abb. 2.1.6-6 Mögliche Bahnkurven im Zentralfeld; verbotene Gebiete *schraffiert*

Gesamtenergie entscheidend (Abb. 2.1.6-5). Bei der kleinstmöglichen Energie (noch kleinere sind nicht realisierbar) ist nur ein fester Wert R des Radius möglich: Das Teilchen läuft auf der *Kreisbahn* mit dem Radius R. Bei größeren Energien entstehen zwei Umkehrpunkte r_{min} und r_{max} für die radiale Bewegung. Die Bewegung des Teilchens ist auf dem *Kreisring* $r_{min} \leqq r \leqq r_{max}$ beschränkt. Bei noch größeren Energien entfällt schließlich der äußere Umkehrpunkt. Das Teilchen kommt etwa aus dem Unendlichen, nähert sich dem Zentrum bis zu einem minimalen Abstand r_0 und entfernt sich dann wieder. Das Teilchen wird abgelenkt, gestreut. Den drei Fällen entsprechen also verschiedene Bahntypen (Abb. 2.1.6-6).

Den gebundenen Bewegungen entsprechen, mit Ausnahme der immer möglichen Kreisbahnen, im allgemeinen keine geschlossenen Bahnen. Bei zwei einfachen und wichtigen Bindungspotentialen sind die Bahnen immer geschlossen: Beim *dreidimensionalen isotropen harmonischen Oszillator* mit dem Potential $V \sim r^2$ entstehen Ellipsen mit dem Mittelpunkt im Kraftzentrum, das ist also die Ruhelage des Oszillators (s. a. 2.1.4.3.).

Beim Potential $V = -\text{const}/r$, das mit verschiedenen Werten der Konstante in der Himmelsmechanik die gravitative Anziehung von Massen und in der Atommechanik die Coulombsche Anziehung zwischen Atomkern und Elektronen erfaßt, entstehen Ellipsenbahnen mit einem Brennpunkt im Kraftzentrum.

2.1.6.4. Zweiteilchenproblem. Das Kraftzentrum des auf ein Teilchen wirkenden Zentralfelds (↑ 2.1.6.3.) wird meist durch ein anderes Teilchen gebildet. Es handelt sich dann eigentlich nicht um die Bewegung eines Teilchens um ein festes Zentrum, sondern um die Bewegung zweier Teilchen „umeinander". Das Zentralpotential $V(r)$ ist in Wirklichkeit das Potential $V_W(r)$ der gegenseitigen Wechselwirkung (↑ 2.1.5.5.) beider Teilchen, wobei $r = |\boldsymbol{r}_1 - \boldsymbol{r}_2|$ der Abstand beider Teilchen ist.

Wenn das eine der beiden Teilchen eine sehr viel größere Masse als das andere hat, wie beim Vergleich Sonne – Planet oder Atomkern – Elektron, so kann seine Bewegung genähert vernachlässigt werden. In dieser Näherung eines ruhenden „Zentralteilchens", das als Kraftzentrum wirkt, bewegt sich das leichtere Teilchen genau so wie eben diskutiert (↑ 2.1.6.3.).

Die Näherung eines schweren ruhenden Teilchens ist jedoch unnötig. Das Bewegungsproblem zweier nur miteinander wechselwirkender Teilchen, auf die also nicht noch zusätzlich „äußere" Kräfte einwirken, kann exakt auf das Zentralfeldproblem zurückgeführt werden. Die Lage beider Teilchen ist eindeutig bestimmt, wenn die Lage ihres Schwerpunkts und ihr

Verbindungsvektor $r \equiv r_1 - r_2$ bekannt ist (Abb. 2.1.6-7). Der Schwerpunkt s teilt die Verbindungslinie beider Teilchen im umgekehrten Verhältnis ihrer Massen, er liegt also näher

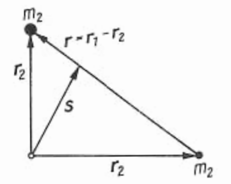


Abb. 2.1.6-7 Schwerpunkts- und Verbindungsvektor zweier Teilchen

bei dem schwereren Teilchen und bei gleichen Massen genau in der Mitte. Da voraussetzungsgemäß keine äußeren Kräfte wirken, bewegt sich der Schwerpunkt mit konstanter Geschwindigkeit. Der Verbindungsvektor r ändert sich mit der Zeit genauso wie der Ortsvektor eines fiktiven Teilchens, das sich mit der Masse $\mu = \frac{m_1 m_2}{m_1 + m_2}$ bzw. $\frac{1}{\mu} + \frac{1}{m_1} + \frac{1}{m_2}$ in dem Zentralfeld $V_W(r)$ bewegt. Das fiktive „Ersatzteilchen" unterliegt also der gleichen Kraft, wie sie von einem der beiden wirklichen Teilchen auf das jeweils andere ausgeübt wird; μ heißt *reduzierte Masse* des Zweiteilchensystems. Sie ist immer kleiner als die kleinere der beiden Teilchenmassen. Diese verringerte „effektive Trägheit" ist anschaulich verständlich. Der Vektor r kann sich leichter ändern, wenn sich beide Teilchen umeinander bewegen, als wenn sich nur eins um ein festes Zentrum bewegt. Falls eine der Massen m_1 und m_2 sehr viel größer als die andere ist, ist die reduzierte Masse μ nur sehr wenig von der kleineren der beiden Teilchenmassen verschieden. Für zwei Teilchen gleicher Masse ist μ gleich der halben Masse eines Teilchens.

Man kann sagen, daß bei Änderung des Vektors r, der die Relativbewegung beider Teilchen beschreibt, sich die beiden Teilchen um ihren gemeinsamen Schwerpunkt bewegen. Der Konstruktion des Schwerpunkts entsprechend verhalten sich dabei die Abstände vom Schwerpunkt und die im Schwerpunktsystem gemessenen Geschwindigkeiten beider Teilchen stets umgekehrt wie ihre Massen.

2.1.6.5. Beschleunigung und Ablenkung im elektrischen Feld. In einem elektromagnetischen Feld (↑ 2.2.3.) wirkt auf ein geladenes Teilchen mit der Ladung q die Kraft $F = q(E + v \times B)$, wobei E das elektrische, B das magnetische Feld und v die Teilchengeschwindigkeit sind. Bei schnell bewegten Teilchen ist die relativistische Bewegungsgleichung (↑ 2.1.5.1.) zu benutzen.

Statische elektrische Felder werden zur Beschleunigung (z. B. in Teilchenbeschleunigern, ↑ 3.1.1.) und Ablenkung (z. B. in Elektronengeräten aller Art ↑ 1.1.3.), geladener Teilchen benutzt. Im einfachsten Falle durchlaufen dabei die Teilchen ein homogenes elektrisches Feld. Wird das Teilchen mit vernachlässigbarer Anfangsgeschwindigkeit in das Feld gebracht, so fliegt es in Feldrichtung. Durchläuft es eine Strecke l im Feld, so leistet die elektrische Kraft die Arbeit $A = Fl = qEl = qU$, wobei U die elektrische Potential- bzw. Spannungsdifferenz ist. Diese Arbeit dient zur Vergrößerung der kinetischen Energie des Teilchens (↑ 2.1.5.4.). Da die kinetische Anfangsenergie als Null angenommen wurde, ist also qU die *kinetische Energie* nach Durchlaufen des Felds. Im nichtrelativistischen Grenzfall gilt einfach $mv^2/2 = qEl$ oder $v = \sqrt{2qEl/m}$. Für ein relativistisches Teilchen (↑ 2.1.3.5.) ist die kinetische Energie in der Form $mc^2[(1 - v^2/c^2)^{-1/2} - 1]$ darzustellen. Daraus folgt durch Auflösen $\left(\frac{v}{c}\right)^2 = 1 - \frac{1}{(1+u)^2} = \frac{u^2 + 2u}{(1+u)^2}$ mit $u = \frac{qEl}{mc^2}$. Für $u \ll 1$ gilt die nichtrelativistische Näherung; für $u \gg 1$ geht v gegen c (Abb. 2.1.6-8).

Wird ein geladenes Teilchen *senkrecht zur Feldrichtung* mit der Anfangsgeschwindigkeit v_0 in ein homogenes elektrisches Feld hineingeschossen (Abb. 2.1.6-9), so wird es im Feld beschleunigt und abgelenkt. Der Impuls in x-Richtung bleibt konstant; er stimmt also ständig mit seinem Anfangswert p_0 überein. Der Impuls in y-Richtung genügt der Gleichung $dp_y/dt = qE$. Da der Anfangswert von p_y verschwindet, wächst also der Impuls $p_y = qEt$ in Feldrichtung linear mit der Zeit t, die das Teilchen bereits im Feld fliegt. Der *Ablenkwinkel* ist aus $\tan\alpha = p_y/p_x = qEt/p_0$ zu bestimmen. Im nichtrelativistischen Grenzfall folgt aus $p_x = p_0 =$ const. sofort, daß auch die

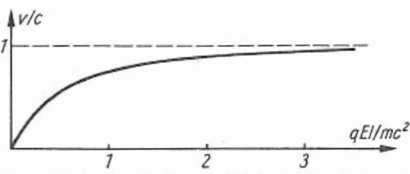


Abb. 2.1.6-8 Endgeschwindigkeit bei Beschleunigung im homogenen elektrischen Längsfeld

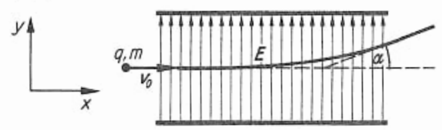


Abb. 2.1.6-9 Ablenkung durch ein homogenes elektrisches Querfeld (für eine positive Ladung gezeichnet)

Geschwindigkeit $v_x = p_x/m = v_0$ in x-Richtung ihren Anfangswert v_0 beibehält. Damit ist die Bestimmung der *Flugzeit* $t = l/v_0$ elementar, und $\tan \alpha = qEl/mv_0^2$ kann unmittelbar durch die (im Experiment leicht kontrollierbare) Länge der Ablenkstrecke ausgedrückt werden. Auch die Bahnkurve kann leicht ermittelt werden: Genau wie beim Wurf im Schwerefeld ist es eine *Parabelbahn.*

Im relativistischen Fall folgt aus der Konstanz von p_x nicht die von v_x, denn die beide Größen verbindende geschwindigkeitsabhängige Masse $m(v)$ wächst wegen der Beschleunigung des Teilchens, so daß v_x abnimmt. Die Flugzeit verlängert sich also und mit ihr auch der Ablenkwinkel. Die Rechnung liefert für die Flugzeit $v_0t/l = (e^{\gamma} - e^{-\gamma})/2\gamma$ mit $\gamma = qEl/p_0c$.

2.1.6.6. Kreisbewegung im magnetischen Feld. Statische Magnetfelder werden häufig zur Ablenkung geladener Teilchen benutzt, z. B. in Teilchenbeschleunigern und Elektronenmikroskopen. Da die magnetische Kraft keine Arbeit verrichtet (↑ 2.1.5.4.), bleibt die Bahngeschwindigkeit des Teilchens unverändert. Daher bleibt auch die geschwindigkeitsabhängige Masse $m(v)$ bei einer jeden Bewegung konstant, so daß das relativistische Bewegungsproblem nicht komplizierter als das nichtrelativistische ist. Wird ein geladenes Teilchen mit der Geschwindigkeit v, also mit dem Impuls $p = m(v)\, v$ senkrecht zu den Feldlinien eines homogenen Magnetfelds B eingeschossen, so beschreibt es unter dem Einfluß der senkrecht auf der Bahnkurve stehenden magnetischen Kraft $q\boldsymbol{v} \times \boldsymbol{B}$, deren Betrag wegen $\boldsymbol{v} \perp \boldsymbol{B}$ einfach qvB ist, eine *Kreisbahn* (Abb. 2.1.6-10). Die zum Kreismittelpunkt gerichtete Beschleunigung der Kreisbewegung (↑ 2.1.4.2.) hat den Betrag $a = v^2/R$, R ist der Bahnradius. Die Bewegungsgleichung fordert $m(v)\, b = qvB$, also $\frac{v}{R} = \frac{qB}{m} \sqrt{1 - v^2/c^2}$. Das Verhältnis von Geschwindigkeit und Bahnradius ist gleich der Winkelgeschwindigkeit (↑ 2.1.4.2.). Diese wird

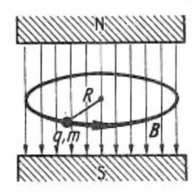


Abb. 2.1.6-10 Kreisbahn im Magnetfeld (für eine positive Ladung gezeichnet)

für ein geladenes Teilchen als *Zyklotronfrequenz* bezeichnet. Für kleine Geschwindigkeiten $v \ll c$ ist sie konstant, für große nimmt sie entsprechend dem Faktor $\sqrt{1 - v^2/c^2}$ ab. In der Form $p/R = qB/m$ wird deutlich, daß bei gegebenem Magnetfeld der *Impuls proportional zum Bahnradius* ist. Darauf beruhen die Impulsmessungen von geladenen Elementarteilchen bei der *Auswertung von Blasenkammeraufnahmen* (↑ 3.2.1.2.2.).

2.2. Elektrische und magnetische Felder

2.2.1. Feldbegriff

Neben Teilchen existieren in der Natur auch nicht scharf lokalisierte Phänomene, die ganze Raumgebiete erfassen. Sie werden durch Felder beschrieben. In ihnen hat die durch sie dargestellte physikalische Größe, die Feldgröße, an jedem Ort einen bestimmten Wert, und dieser Wert kann sich im Laufe der Zeit ändern.

Auf diese Weise kann der Zustand einer Substanz, beispielsweise der Temperaturverlauf im Inneren eines ungleichmäßig erwärmten Körpers, beschrieben werden. Zu jedem Ort $\boldsymbol{r}$ im Körper gehört dann eine Temperatur $T(\boldsymbol{r})$, die

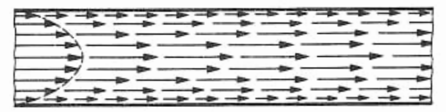


Abb. 2.2.1-1 Querschnitt des Geschwindigkeitsfelds einer Rohrströmung

hier die skalare Feldgröße darstellt. Das Temperaturfeld ist im allgemeinen zeitabhängig, da sich die Temperaturen ändern können. Einen Überblick über den Feldverlauf zu einer bestimmten Zeit erhält man, indem man die Flächen konstanter Temperaturen angibt.

Das vektorielle Geschwindigkeitsfeld $\boldsymbol{v}(\boldsymbol{r})$ dagegen kennzeichnet einen Strömungsvorgang; diese Feldgröße gibt an, mit welcher Geschwindigkeit $\boldsymbol{v}$ sich die Substanz am betrachteten Ort bewegt (Abb. 2.2.1-1). Dieses Feld ändert sich im allgemeinen ebenfalls mit der Zeit und ist nur für einen stationären Strömungsvorgang zeitunabhängig.

Ein zeitunabhängiges Feld wird als *statisches Feld* bezeichnet. Zeitabhängige Felder können typische Ausbreitungsvorgänge zeigen, die man *Wellen* nennt (↑ 2.3.). Bei geeigneter räumlicher Begrenzung sind *Eigenschwingungen des Felds* möglich (↑ 2.3.3.), die oft auch als *stehende Wellen* bezeichnet werden. Langsam zeitabhängige Felder, bei denen die Ausbreitungsgeschwindigkeit des Felds noch nicht wichtig ist, heißen *quasistationär*. Quasistationäre elektrische und magnetische Felder sind für elektrische Wechselströme typisch.

2.2.2. Ladung und Strom

2.2.2.1. Elektrische Ladung. Elektrische Ladungen sind definiert durch die beiden Eigen-

schaften, elektrische Felder zu erzeugen und in elektrischen Feldern Kräfte zu erfahren. Das Experiment zeigt die Existenz von Ladungen mit verschiedenen Vorzeichen. Die meisten Elementarteilchen tragen eine positive oder negative *elektrische Elementarladung* (↑ 1.1.2.1.). Diese hat den Wert $e = 1{,}602 \cdot 10^{-19}$ As. Die Zuordnung der Vorzeichen ist an sich willkürlich. Eine gleichzeitige Umbenennung aller positiven Ladungen in negative und umgekehrt würde an den beobachtbaren Folgerungen aus der Theorie nichts ändern.

Für elektrische Ladungen gilt der *Ladungserhaltungssatz*. Solange keine Erzeugung, Vernichtung bzw. Umwandlung von Teilchen stattfindet, wird diese Erhaltung einfach dadurch gewährleistet, daß jedes Teilchen seine Ladung unverändert beibehält. Bei Umwandlungsprozessen zwischen Elementarteilchen entstehen bzw. verschwinden stets gleich viel positive und negative Elementarladungen. In einem beliebigen begrenzten Raumgebiet kann sich die elektrische Gesamtladung nur dadurch verändern, daß Ladungen aus diesem Gebiet heraus oder andere hinein wandern, daß also Ströme durch die Oberfläche des Gebiets fließen. Häufig wird die *räumliche Verteilung* der elektrischen Ladungen nicht im Bilde einzelner geladener Teilchen beschrieben, sondern durch eine *Ladungsdichte* ϱ, das ist die Ladung pro Volumen, bestimmt für ein hinreichend kleines Volumenelement. $\varrho \Delta V$ ist demnach die in dem Volumenelement ΔV enthaltene Ladung. Eine ungleichmäßige Ladungsverteilung wird durch eine ortsabhängige Dichte $\varrho(\boldsymbol{r})$ beschrieben. Durch die Dichtebeschreibung wird über die Positionen einzelner geladener Teilchen gemittelt. Das ist zweckmäßig bzw. notwendig für alle makroskopischen Erscheinungen, an denen ungeheuer viel Elementarladungen beteiligt sind; aber auch dann, wenn wie bei den Elektronen im Atom eine genaue Lokalisierung durch die quantenmechanische Unbestimmtheitsrelation (↑ 2.5.1.7.) verhindert wird.

2.2.2.2. Elektrischer Strom. Elektrische Ströme entstehen durch Bewegung elektrischer Ladungen. Unter dem mittleren elektrischen Strom $I = \Delta Q/\Delta t$ durch eine Fläche A versteht man das Verhältnis der durch A hindurch transportierten Ladungsmenge ΔQ zu der Zeitdauer Δt, in der dieser Transport erfolgt. Dabei muß eine Transportrichtung an sich willkürlich als positiv definiert werden. Ein Ladungstransport in entgegengesetzter Richtung trägt mit umgekehrtem Vorzeichen zum Strom bei; das Vorzeichen der Ladung ist ebenfalls zu beachten. Als Beispiel werden vier geladene Teilchen betrachtet (Abb. 2.2.2-1). Zum Strom tragen diejenigen bei, deren Bahnen in der betrachteten Zeit Δt die Fläche A kreuzen. Das Teilchen mit der Ladung q_2 transportiert positive Ladung von links nach rechts, q_3 dagegen negative Ladung von rechts nach links. Insgesamt fließt der Strom $[q_2 + q_3]/\Delta t$ von links nach rechts. Wählt man die Zeitdauer Δt so klein, daß nur wenige geladene Teilchen die Fläche passieren, so wird der Strom von einem Zeitintervall zum nächsten starke *Schwankungen* zeigen (Schroteffekt). Makroskopische Ströme entstehen durch Beteiligung ungeheuer vieler Elementarladungen (Elektronen in Metallen). Das Intervall Δt kann dann sehr klein sein, ohne daß die erwähnten Schwankungen merklich sind.

Eine detaillierte Beschreibung des Ladungstransports wird durch Einführung der elektrischen *Stromdichte* $\boldsymbol{j}$ erreicht. Deren Betrag ist als Quotient aus Strom ΔI und einem zur Strömungsrichtung senkrechten Flächenelement ΔA definiert. Strömen mit der Ladungsdichte ϱ im Raum verteilte Ladungen mit einer einheitlichen Strömungsgeschwindigkeit $\boldsymbol{v}$, so gilt $\boldsymbol{j} = \varrho \boldsymbol{v}$.

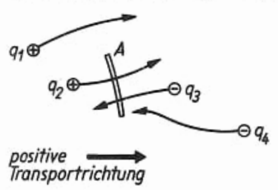


Abb. 2.2.2-1 Zum Strombegriff; $q_1 \dots q_4$ bewegte Ladungen

2.2.2.3. Stromeinheit und Ladungseinheit. Wegen des Zusammenhangs $I = \Delta Q/\Delta t$ zwischen Strom I und Ladung ΔQ genügt es, für eine dieser beiden Größen eine Einheit festzulegen. Eine Ladungseinheit könnte über das Kraftgesetz zwischen elektrischen Ladungen definiert werden. Meßtechnisch wesentlich genauer ist aber die Festsetzung der *Stromeinheit* durch Messung der magnetischen Kraft zwischen zwei parallelen geraden Drähten, durch die der gleiche Strom fließt (↑ 2.2.3.5.). Definitionsgemäß fließt ein Strom von der Stromstärke 1 Ampere = 1 A, wenn bei einem Drahtabstand von 1 m auf ein Drahtstück der Länge 1 m eine Kraft von $2 \cdot 10^{-7}$ N $= 2 \cdot 10^{-7}$ kgms^{-2} wirkt. Ströme von 1 bis 10 A sind typisch für elektrische Haushaltgeräte. Dagegen werden Industrieverbraucher häufig für Stromstärken im Bereich von 10^5 bis 10^6 A abgesichert.

Die Ladungseinheit ist eine Amperesekunde = 1 As, die auch als 1 Coulomb = 1 C bezeichnet wird. Bei elektrostatischen Anordnungen treten stets sehr viel kleinere Ladungen auf. Würden sich zwei Ladungen von je 1 As im Abstand von 1 km (!) voneinander befinden, so entstünde eine Kraft von $9 \cdot 10^3$ N, $\approx 10^4$ N, das entspricht dem Gewicht eines Körpers von etwa 1 t $= 10^3$ kg.

2.2.2.4. Elektrisches Dipolmoment. Bei zwei häufig auftretenden Fragestellungen kommt es nicht auf die genaue Verteilung (Anordnung) der elektrischen Ladungen innerhalb eines Systems an: bei der Bestimmung des von den Ladungen in großem Abstand erzeugten elektrischen Felds und bei der Reaktion des Systems auf ein im Bereich der Ladungsverteilung nahezu homogenes äußeres elektrisches Feld. Ist die *Gesamtladung* von Null verschieden, so bestimmt sie das Verhalten des Systems, das sich dann fast so verhält, als wäre die Gesamtladung in einem Punkt konzentriert.

Bei vielen wichtigen Systemen ist die Gesamtladung Null, da sie „gleich viele" positive und negative Ladungen enthalten. Beispielsweise gilt dies für (nicht ionisierte) Atome und Moleküle. Solche Systeme können dennoch Felder erzeugen und mit Feldern in Wechselwirkung treten. Sie tun dies insbesondere dann, wenn die positiven und negativen Ladungen im Mittel gegeneinander verschoben sind, so daß die „Schwerpunkte" der Ladungen beider Vorzeichen nicht zusammenfallen. Die Kenngröße einer solchen Anordnung ist der Vektor $\boldsymbol{m}_{el}$ des elektrischen Dipolmoments. Er zeigt vom negativen zum positiven Ladungsschwerpunkt; in diesem Sinne trägt ein elektrischer Dipol „hinten" negative und „vorn" positive Ladungen. $m_{el} = Q_+ l$ bestimmt die Stärke des Dipolmoments. Dabei ist Q_+ die Summe aller positiven Ladungen und l der Abstand der beiden Ladungsschwerpunkte. Wegen der verschwindenden Gesamtladung gilt $Q_- = -Q_+$ für die Summe der negativen Ladungen.

2.2.2.5. Magnetisches Dipolmoment. Ein stromdurchflossener Ring verhält sich hinsichtlich der beiden Fragestellungen „Felderzeugung" und „Verhalten im äußeren Feld" (↑ 2.2.2.4.) bei Betrachtung magnetischer Felder völlig analog wie ein elektrischer Dipol. Der einfachste Fall ist ein ebener *Ringstrom*. Sein magnetisches Dipolmoment $\boldsymbol{m}_{mg}$ steht senkrecht auf der Ebene des Ringstroms, und zwar so, daß $\boldsymbol{m}_{mg}$ zusammen mit der Stromrichtung eine Rechtsschraube bildet. Der Betrag des Dipolmoments ist $m_{mg} = IA$, wobei I die Stromstärke und A die vom Stromring eingeschlossene Fläche ist. Das Dipolmoment kann für jede Stromverteilung definiert werden, die man sich aus Ringströmen aufgebaut denken kann, z. B. für eine rotierende geladene Kugel.

Oft wird aus historischen Gründen das magnetische Dipolmoment mit einem zusätzlichen Faktor μ_0 versehen (↑ 2.2.3.4.).

Ein Ringstrom wird durch umlaufende geladene Teilchen erzeugt. Da diese außer Ladung auch Masse tragen, besitzt das System einen inneren Drehimpuls, also eine *Spin* $\boldsymbol{s}$ (↑ 2.1.3.9.). Entsteht der Strom nur durch die (Bahn-) Bewegung von Ladungsträgern einer Sorte, so sind magnetisches Moment und Gesamtdrehimpuls unmittelbar proportional. Es gilt $\boldsymbol{m}_{mg} = (q/2m)\,\boldsymbol{s}$, wobei q und m Ladung und Masse der einzelnen Teilchen bezeichnen. Auch die Spins von Elementarteilchen sind mit *magnetischen Eigenmomenten* dieser Teilchen verbunden. Der Proportionalitätsfaktor g in $\boldsymbol{m}_s = g(q/2m)\,\boldsymbol{s}$ heißt *gyromagnetischer Faktor* oder kurz g-Faktor und ist im allgemeinen von 1 verschieden. Für Elektronen ist mit großer Genauigkeit $g = 2$, das mit dem Elektronenspin verbundene magnetische Moment ist also doppelt so groß wie für einen Bahndrehimpuls gleicher Größe.

Da der Spin zusammengesetzter Teilchen, wie etwa der von Atomkernen, Atomen, Molekülen, sich aus Bahn- und Spindrehimpulsen von Elementarteilchen zusammensetzt, ist auch der g-Faktor solcher Systeme von 1 und damit auch von Kern zu Kern, Atom zu Atom verschieden (s. a. 4.2.4.5.).

2.2.3. Statische elektrische und magnetische Felder

Aufgrund der Erfahrungen bei zeitabhängigen elektromagnetischen Feldern (↑ 2.2.4.) werden auch die statischen Felder als eigenständige physikalische Realitäten betrachtet. Sie sind also nicht nur formales Hilfsmittel zur Beschreibung elektrischer und magnetischer Kräfte. Während man vor Aufstellung der speziellen Relativitätstheorie den Äther, d. h. eine besondere, alles durchdringende Substanz, als Träger der elektrischen und magnetischen Felder annahm, wurde danach dieses Konzept als undurchführbar fallen gelassen.



Abb. 2.2.3-1 Elektrisches Feld außerhalb einer positiv geladenen Kugel

2.2.3.1. Elektrische Felder. Elektrische Felder sind Vektorfelder; der ortsabhängige Feldvektor wird mit $\boldsymbol{E}(\boldsymbol{r})$ bezeichnet. Statische elektrische Felder werden von ruhenden elektrischen Ladungen, d. h. zeitunabhängigen Ladungsdichten, erzeugt. Eine *kugelsymmetrische*

Ladungsverteilung erzeugt ein radial gerichtetes, ebenfalls kugelsymmetrisches Feld (Abb. 2.2.3-1). Außerhalb einer Kugel mit der Gesamtladung Q entsteht ein elektrisches Feld, $E = Q/4\pi\varepsilon_0 r^2 \approx$ mit $1/4\pi\varepsilon_0 \approx 9 \cdot 10^9$ Vm(As)$^{-1}$ dessen Stärke mit dem Abstand r vom Kugelmittelpunkt abnimmt. Für $Q > 0$ ist das Feld nach außen, für $Q < 0$ nach innen gerichtet. Die Einheit der Feldstärke ist Vm^{-1}. Die *Spannungseinheit* (↑ 2.2.3.2.) 1 Volt = 1 V ist eine abgeleitete Einheit, sie hängt über 1 VAs = 1 Nm = 1 J mit der mechanischen Energieeinheit (↑ 2.2.5 3.) zusammen. Die Konstante ε_0 kann durch Messung der elektrischen Kraft (↑ 2.2.3.3.) zwischen zwei Kondensatorplatten bestimmt werden. Gute Isolatoren besitzen Durchschlagsfeldstärken bis zu 10^7 Vm^{-1}. An metallischen Spitzen im Vakuum können Feldstärken bis zu 10^9 Vm^{-1} erzeugt werden.

Für das idealisierte Modell einer Punktladung ist der zu $1/r^2$ proportionale Feldverlauf bis zu beliebig kleinen Abständen r gültig. Die Feldstärke wird in der Nähe der Punktladung beliebig groß. Für eine Kugel mit endlichem Radius wird der singuläre Feldverlauf „abgeschnitten", so daß die Feldstärke im Inneren der Kugel endlich bleibt. Der genaue Verlauf des Innenfelds hängt von den Details der Ladungsverteilung ab.

2.2.3.2. Spannung und Potential. Das negative Wegintegral der elektrischen Feldstärke $U_{12} = -\int_{r_1}^{r_2} E\,dr$, das so wie das analoge Arbeitsintegral (↑ 2.1.5.5.) gebildet wird, heißt *elektrische Spannung*. Die Spannung U_{12} hängt bei zeitabhängigen Feldern von dem gewählten Weg zwischen den beiden Punkten 1 und 2 ab. U_{12} ist bei statischen elektrischen Feldern unabhängig vom Integrationsweg. Bei festgehaltenem Anfangspunkt 1 hängt U_{12} dann nur noch vom Endpunkt 2 ab. Da dieser beliebig wählbar ist, wird jedem Raumpunkt ein bestimmter Spannungswert als Spannungsdifferenz zum willkürlichen Anfangspunkt zugeordnet. Dieses Spannungsfeld heißt *elektrisches Potential* φ. φ nimmt ab, wenn man in Feldrichtung weitergeht. Bei einem homogenen elektrischen Feld gilt $|U_{12}| = EL$ für zwei Punkte, die in Feldrichtung um die Strecke L auseinanderliegen.

2.2.3.3. Kraft und Drehmoment im elektrischen Feld. Auf eine elektrische Ladung q wirkt im elektrischen Feld eine Kraft $\boldsymbol{F} = q\boldsymbol{E}$, die proportional zur Feldstärke $\boldsymbol{E}$ am Ort der Ladung ist. Dabei wird vorausgesetzt, daß die Ladung in einem Raumbereich lokalisiert ist, in dem das Feld $\boldsymbol{E}$ als konstant angesehen werden kann. Die Kraft zeigt in Feldrichtung bei einer positiven, in entgegengesetzter Richtung bei einer negativen Ladung. Das von q selbst erzeugte Eigenfeld trägt zur Kraft nicht bei.

Stehen zwei Ladungen q_1 und q_2 miteinander in Wechselwirkung, so erzeugt jeweils eine von ihnen das auf die andere wirkende elektrische Feld. Ist r der Abstand beider Ladungen, so erzeugt q_1 bei q_2 das Feld $E = q_1/4\pi\varepsilon_0 r^2$, so daß $F = q_1 q_2/4\pi\varepsilon_0 r^2$ die Kraft auf q_2 ist. Bis auf die Richtung erhält man die gleiche Kraft auf q_1 in dem von q_2 erzeugten Feld. Die elektrostatische Wechselwirkung genügt also dem *Gegenwirkungsprinzip* (↑ 2.1.5.1.). Das eben abgeleitete *Kraftgesetz* ist das Coulombgesetz, die Wechselwirkung wird oft als *Coulombwechselwirkung* bezeichnet. Bei gleichem Ladungsvorzeichen ergibt sich Abstoßung, sonst Anziehung.

Da qE die Kraft auf eine Ladung q darstellt, ist $qU_{12} = A_{12}$ die von dem elektrischen Feld verrichtete Arbeit; U_{12} ist die Spannungsdifferenz (↑ 2.2.3.2.). Aus der Existenz eines elektrischen Potentials $\varphi(r)$ für elektrostatische Felder folgt so die Existenz des Potentials $V(r) = q\varphi(r)$ (↑ 2.1.5.6.) für die elektrische Kraft. Elektrostatische Kräfte sind also konservative Kräfte.

Durchläuft ein Teilchen mit der Ladung q in einem beschleunigenden Feld eine Spannungsdifferenz ΔU (s. a. 2.1.6.5.), so erhöht sich seine kinetische Energie um $q\,\Delta U$. Bei Strömen in Drähten wird diese kinetische Energie der Elektronen durch Stöße mit den Atomen in *Joulesche Wärme* überführt. An einem Draht, an einem Widerstand, liegt dann die Spannung 1 V, wenn bei einem Strom von 1 A in 1 s die Energie 1 J = 1 Ws in Wärme übergeführt wird. Durchläuft ein Teilchen mit einer Elementarladung im Vakuum eine Spannungsdifferenz von 1 V, so gewinnt es die Energie 1 Elektronenvolt = 1 eV = $1{,}602 \cdot 10^{-19}$ VAs = $1{,}602 \cdot 10^{-19}$ J. Die *Energieeinheit Elektronenvolt* ist in der Atomphysik sehr zweckmäßig, da die auftretenden Energien von dieser Größenordnung sind.

Ein elektrischer Dipol (↑ 2.2.2.4.) erfährt im elektrischen Feld ein *Drehmoment* $\boldsymbol{M} = \boldsymbol{m}_{el} \times \boldsymbol{E}$, das bestrebt ist, den Dipol in Feldrichtung zu drehen. Der Dipol verhält sich wie ein System mit der potentiellen Energie $E_p = -\boldsymbol{m}_{el}\boldsymbol{E}$. Diese hat ihren kleinsten Wert bei Parallelstellung von Dipolmoment und äußerem Feld; das Drehmoment verschwindet dann. Ohne zusätzliche Einflüsse (z. B. Reibung) schwingt der Dipol im Feld um diese Gleichgewichtslage.

2.2.3.4. Magnetische Felder. Statische Magnetfelder werden von stationären elektrischen Strömen erzeugt. Ein *gerader stromdurchflossener Draht* ist von kreisförmigen magnetischen Feldlinien umschlossen (Abb. 2.2.3-2). Der Betrag B des magnetischen Felds ist proportional zur Stromstärke I und nimmt außerhalb des Drahts entsprechend $B = \mu_0 J/2\pi r$ mit

$\mu_0 = 4\pi \cdot 10^{-7}$ Vs(Am)$^{-1}$ mit wachsendem Abstand r von der Drahtachse ab. Feld- und Stromrichtung bilden eine *Rechtsschraube.* Die *Einheit des magnetischen B-Felds,* das auch als magnetische Induktion bezeichnet wird, ist 1 Vsm^{-2} = 1 Tesla = 1 T. Die Voltsekunde

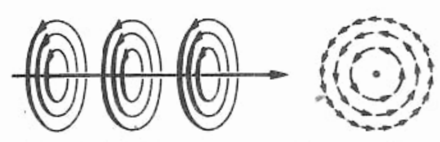


Abb. 2.2.3-2 Magnetisches Feld eines stromdurchflossenen Drahts

ist die *Einheit des magnetischen Flusses* (↑ 2.2.3.6.), für sie ist die Bezeichnung 1 Vs = 1 Weber = 1 Wb möglich. Es gilt also 1 T = 1 Wbm^{-2}. Eine veraltete, aber noch anzutreffende Einheit ist 1 Gauß = 1 G, zur Umrechnung dient 1 T = 10^4 G. Die Konstante μ_0 kann durch Messung der magnetischen Kraft (↑ 2.2.3.5.) zwischen stromdurchflossenen Drähten bestimmt werden; ihr Zahlenwert ist jedoch kein gemessener Wert, sondern wird im Zusammenhang mit der Definition der Stromeinheit 1 Ampere festgelegt. Im Inneren einer *stromdurchflossenen Spule,* deren Länge sehr groß im Vergleich zu ihrem Durchmesser ist, hat das Magnetfeld überall gleiche Richtung und Stärke; es ist homogen. Seine Stärke ist durch den Strom I in der Spule, ihre Länge L und ihre Windungszahl N bestimmt zu $B = \mu_0 NI/L$.

Sowohl für die magnetische Kraft als auch für die Induktionswirkungen (↑ 2.2.4.1.) ist die Feldgröße $\boldsymbol{B}$ maßgebend. Dennoch wird aus historischen Gründen, weil nämlich die ersten magnetischen Erfahrungen und die entsprechenden Begriffsbildungen mit Permanentmagneten gewonnen wurden, während erst später von OERSTEDT die grundlegenderen magnetischen Wirkungen elektrischer Ströme entdeckt wurden, auch zur Beschreibung von Feldern im Vakuum noch eine zweite magnetische Feldgröße, die sogenannte *magnetische Feldstärke* $\boldsymbol{H}$, benutzt. Während $\boldsymbol{H}$ und $\boldsymbol{B}$ sich im Inneren magnetisierter Substanzen wesentlich unterscheiden können, hängen sie im Vakuum (und in sehr guter Näherung auch in Luft) einfach durch $\boldsymbol{B} = \mu_0 \boldsymbol{H}$ miteinander zusammen, so daß eine von beiden Größen natürlich entbehrlich wäre. Es ist daher bei solchen Problemen auch belanglos, ob ein magnetisches Feld durch die magnetische Induktion $\boldsymbol{B}$ oder durch die magnetische Feldstärke $\boldsymbol{H}$ beschrieben wird. In einigen Gebieten der Physik wird jeweils eine der beiden Feldgrößen und die zugehörigen Einheiten bevorzugt benutzt.

Die Einheit der magnetischen Feldstärke ist Am^{-1}. Eine veraltete Einheit ist 1 Oerstedt = 1 Oe = $(10^3/4\pi)$ Am^{-1}. Durch B = 1 G und H = 1 Oe wird demnach dasselbe Magnetfeld beschrieben. Mit $4\pi \approx 10$ entspricht ein H-Feld von 100 Am^{-1} einem B-Feld von etwa 1 G bzw. 10^6 Am^{-1} entsprechen etwa 1 T.

In den Spulen und Transformatoren der Elektrotechnik liegen die Maximalwerte der magnetischen Feldstärke meist zwischen 1 und 2000 Am^{-1}. Im Luftspalt zwischen den Polschuhen starker Elektromagnete erreicht man Feldstärken von $5 \cdot 10^6$ Am^{-1} und mit supraleitenden Magnetspulen von mehr als 10^7 Am^{-1}, also B-Felder von einigen 10 T.

2.2.3.5. Kraft und Drehmoment im magnetischen Feld. Auf eine mit der Geschwindigkeit $\boldsymbol{v}$ *bewegte elektrische Ladung q* wirkt in einem magnetischen Feld mit der Induktion $\boldsymbol{B}$ eine Kraft $\boldsymbol{F}_{mg} = q\boldsymbol{v} \times \boldsymbol{B}$ (Abb. 2.2.3-3). Dabei ist vorausgesetzt, daß die Ladung in einem Bereich konzentriert ist, in dem das Magnetfeld als homogen angesehen werden kann. Die Kraft ist senkrecht zur Geschwindigkeit und senkrecht zum Magnetfeld gerichtet. Aus dem Kraftgesetz für eine bewegte Ladung folgt das für die Kraft auf ein kurzes Stück der Länge Δl eines vom Strom I durchflossenen Drahts. Ist ΔQ die Summer aller in diesem Drahtstück zum Stromtransport beitragenden Ladungen und braucht jede die Zeit Δt zum Durchfließen der Strecke Δl, so ist die Stromstärke $I = \Delta Q/\Delta t$ und die Geschwindigkeit der Ladungsträger $v = \Delta l/\Delta t$. Also gilt $v\,\Delta Q = I\,\Delta l$, und die Kraft ist $\boldsymbol{F}_{mg} = J\,\Delta l(\boldsymbol{e} \times \boldsymbol{B})$, wobei $\boldsymbol{e}$ der Einheitsvektor in Stromrichtung und $\boldsymbol{B}$ das Feld beim Drahtelement Δl ist.

Stehen *zwei Ströme* I_1 und I_2 miteinander *in Wechselwirkung,* so erzeugt jeweils einer von ihnen das auf den anderen wirkende magnetische Feld. Fließen die beiden Ströme in langen geraden und zueinander parallelen Drähten mit dem Abstand r (Abb. 2.2.3-4), so ist $B_1 = \mu_0 I_1/2\pi r$ das von I_1 bei I_2 erzeugte Feld, also $F = I_2\,\Delta l B_1 = \mu_0 I_1 I_2/2\pi r$ die Kraft auf ein Drahtstück der Länge Δl. Bis auf die Richtung erhält man die gleiche Kraft auf ein gleich großes Stück des von I_1 durchflossenen Drahts im von I_2 erzeugten Feld. Die magnetischen Wechselwirkungen zwischen stationären Strömen genügen also genau so wie die elektro-

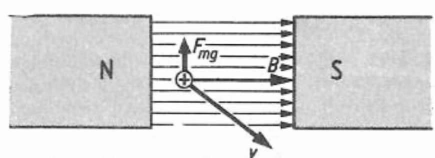


Abb. 2.2.3-3 Magnetische Kraft auf eine bewegte Ladung

statischen Wechselwirkungen dem *Gegenwirkungsprinzip*. Zwei parallele Ströme, die in gleicher Richtung fließen, ziehen sich an, anderenfalls stoßen sie sich ab. Durch Messung der Kraft zwischen parallelen Strömen werden gleichzeitig die Stromeinheit Ampere (↑ 2.2.2.3.) und der Wert von μ_0 durch Definition festgelegt.

Ein magnetischer Dipol (↑ 2.2.2.5.) erfährt im Magnetfeld ein *Drehmoment* $\boldsymbol{M} = \boldsymbol{m}_{\text{mg}} \times \boldsymbol{B}$, völlig analog zum elektrischen Dipol im elektrischen Feld (↑ 2.2.3.3.). Wie dieser besitzt er eine potentielle Energie $-\boldsymbol{m}_{\text{mg}}\boldsymbol{B}$ im äußeren Feld. Da ein magnetischer Dipol im Gegensatz zu einem elektrischen kein „statisches Gebilde" ist, sondern von vornherein eine innere Dynamik, einen Spin (↑ 2.1.3.9., 2.2.2.5.) besitzt, verhält er sich unter dem Einfluß des Drehmoments grundsätzlich anders als ein elektrischer Dipol, nämlich ähnlich wie ein Kreisel.

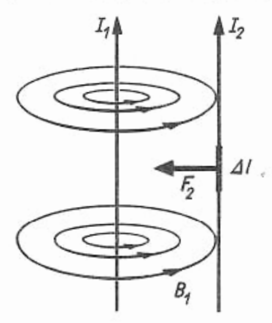


Abb. 2.2.3-4 Zur magnetischen Wechselwirkung zwischen zwei Strömen

Er führt im Magnetfeld eine als *Larmorpräzession* bezeichnete Bewegung aus (s. a. 2.1.5.8.), wobei sich das gesamte System in sich starr mit der Winkelgeschwindigkeit $\omega_L = (m_{\text{mg}}/s)\, B$ um die Magnetfeldrichtung dreht. Spin und magnetisches Moment sind zueinander proportional (↑ 2.2.2.5.). Meist werden s und m_{mg} durch Elektronen erzeugt, die Präzessionsbewegung bildet dann, ebenso wie bei anderen negativen Ladungsträgern, zusammen mit der Feldrichtung eine Rechtsschraube.

2.2.3.6. Magnetischer Fluß. Wenn ein homogenes Magnetfeld der Induktion $\boldsymbol{B}$ eine ebene Fläche der Größe A senkrecht durchsetzt, so nennt man das Produkt $\Phi = BA$ den magnetischen Fluß durch die Fläche. Steht die ebene Fläche schräg zum Feld, so trägt nur die Normalkomponente B_n, das ist die Komponente von $\boldsymbol{B}$ in Richtung der Flächennormalen, zum Fluß bei. Im allgemeinen Fall einer gekrümmten Fläche A und eines inhomogenen Magnetfelds ist der Fluß durch das Flächenintegral $\Phi = \iint \mathrm{d}AB_n$ gegeben.

Der magnetische Fluß hängt nur von der die Fläche begrenzenden Randkurve ab, nicht von der genauen Form dieser Fläche. Er kann als Maß für die Zahl der die Fläche durchstoßenden magnetischen Feldlinien aufgefaßt werden. Die Unabhängigkeit von der genauen Flächenform folgt dann daraus, daß die Feldlinien des (quellfreien) B-Felds nirgends beginnen bzw. enden.

Da die magnetische Induktion in Vsm^{-2} gemessen wird (↑ 2.2.3.4.), ist 1 Voltsekunde = 1 Vs = 1 Weber = 1 Wb die *Einheit des magnetischen Flusses* Φ. Eine veraltete Einheit ist 1 Maxwell = 1 M = 10^{-4} Vs. Es gilt 1 M = 1 G cm² (↑ 2.2.3.4.).

2.2.3.7. Feldenergie. Elektrische und magnetische Felder sind Träger von Energie. Die Feldenergie tritt als qualitativ neue Energieform zu Ruh- und kinetischer Energie von Teilchen (↑ 2.1.3.4., 2.1.3.6.) hinzu. Die Feldenergie wird so definiert, daß bei einem aus bewegten geladenen Teilchen und elektromagnetischen Feldern bestehenden System ganz allgemein die Summe aus den Energien aller Teilchen und der Feldenergie konstant bleibt, obwohl sich die einzelnen Energien ändern. Die Feldenergie ist im gesamten Feld verteilt, so daß jedes kleine Volumenelement ΔV einen bestimmten, zur Größe von ΔV proportionalen Energieinhalt besitzt. Der Quotient aus Energie und Volumen für ein sehr kleines Volumenelement ist die *räumliche Energiedichte* w. Sie setzt sich entsprechend $w = w_{\text{el}} + w_{\text{mg}}$ mit $w_{\text{el}} = \varepsilon_0 E^2/2$ und $w_{\text{mg}} = B^2/2\mu_0$ additiv aus einem elektrischen und einem magnetischen Anteil zusammen. Bei statischen Feldern kann entweder nur elektrische oder nur magnetische Energie vorliegen. Sowohl für elektrische als auch für magnetische Felder ist die Energiedichte proportional zum Quadrat des jeweiligen Felds und damit immer positiv. Sie ist dort groß, die Feldenergie ist also dort konzentriert, wo das Feld stark ist. Die gesamte Feldenergie ergibt sich aus der Energiedichte bei homogenen Feldern durch Multiplikation mit dem Volumen des felderfüllten Raumbereichs, im allgemeinen als Volumenintegral über die Energiedichte.

2.2.3.8. Fern- und Nahwirkung. Die Formeln für die Kraftwirkung zwischen zwei elektrischen Ladungen bzw. zwei Strömen (↑ 2.2.3.3., 2.2.3.5.) enthalten die Felder, die die Kräfte hervorrufen, nicht mehr. Nach den erhaltenen Ergebnissen müßte eine Änderung des Abstands im selben Moment eine Änderung der Kraft bewirken, und zwar bei beliebig großem Abstand der Ladungen bzw. der Ströme. Entsprechendes gilt auch für die Gravitationswechselwirkung zweier Massen nach dem Newtonschen Gravitationsgesetz, in dem im wesentlichen die beiden Massen m_1 und m_2 an die Stelle der Ladungen q_1 und q_2 treten. Die un-

endlich rasche Ausbreitung einer Wirkung nach der *Fernwirkungstheorie* widerspricht aber der Forderung der speziellen Relativitätstheorie, daß es keine größere Signalgeschwindigkeit als die Lichtgeschwindigkeit gibt.

Die Schwierigkeit wird gelöst durch Einführung des elektromagnetischen Felds. Dann erzeugt eine *bewegte* Ladung q_1 ein elektrisches und ein magnetisches Feld. Diese breiten sich mit Lichtgeschwindigkeit aus, gelangen zur entfernten Ladung q_2 und verursachen an ihr Kräfte. Das ist das Schema einer *Nahwirkungstheorie*, in der die dann natürlich zeitabhängigen Felder (↑ 2.2.4.) für das Verständnis der Wechselwirkung unentbehrlich sind. In dem erweiterten Bild braucht zwischen der Kraft auf q_1 und der Kraft auf q_2 nicht der durch actio = reactio (↑ 2.1.5.1.) beschriebene Zusammenhang zu bestehen. Erst im *Grenzfall ruhender Ladungen* entsteht wieder das alte Coulombsche Gesetz. Für Ladungen, die sich mit Geschwindigkeiten bewegen, die sehr klein gegen die Lichtgeschwindigkeit sind, bleibt dieses Gesetz eine sehr gute Näherung. Das trifft insbesondere bei den Bewegungen von Elektronen in Atomen, Molekülen, Festkörpern u. a. zu.

In dem eben diskutierten Grenzfall ist bei elektromagnetischer Wechselwirkung der Teilchen gerade die Voraussetzung für die *Einführung einer potentiellen Energie* (↑ 2.1.5.5.) und damit für die Formulierung eines „mechanischen Energiesatzes", anstelle der allgemeinen Energieerhaltung, erfüllt. Die Freiheitsgrade des mit den Teilchen gekoppelten Systems, hier des elektromagnetischen Felds, lassen sich aus der Beschreibung eliminieren.

2.2.4. Zeitabhängige elektromagnetische Felder

2.2.4.1. Kopplung elektrischer und magnetischer Felder. Während ein elektrostatisches Feld ohne ein „zugehöriges" Magnetfeld und ebenso ein statisches Magnetfeld ohne ein „zugehöriges" elektrisches Feld existieren kann, sind entsprechende zeitabhängige Felder stets miteinander gekoppelt. Dies wird durch den Begriff des *elektromagnetischen Felds* ausgedrückt. Die gegenseitige Kopplung der elektrischen und magnetischen Felder erfolgt durch zwei Effekte:

Ein zeitabhängiges Magnetfeld erzeugt ein elektrisches Wirbelfeld.

Ein zeitabhängiges elektrisches Feld wirkt hinsichtlich der Erzeugung eines Magnetfelds ähnlich wie ein elektrischer Strom.

Der erste Effekt wird durch das *Induktionsgesetz* quantitativ erfaßt. Danach tritt an den Enden einer Leiterschleife eine Spannung auf, wenn sich das die Leiterschleife durchsetzende Magnetfeld zeitlich ändert. Diese induzierte Spannung U_{ind} ist entsprechend $U_{\text{ind}} = -\mathrm{d}\Phi/\mathrm{d}t$ durch die *Änderung des magnetischen Flusses* pro Zeiteinheit bestimmt. Dabei ist $U_{\text{ind}} = \oint \boldsymbol{E}\,\mathrm{d}\boldsymbol{r}$ die Umlaufspannung längs der benutzten Leiterschleife. Diese ist jedoch nur als Hilfsmittel zur Messung zu betrachten, das elektrische Wirbelfeld besteht auch ohne die Leiterschleife. Der Integrationsweg, die Induktionsschleife, umschließt die zur Berechnung von Φ benutzte Fläche. Umlaufsinn und positive Flußrichtung bilden eine Rechtsschraube. Das Induktionsgesetz in der angegebenen Form gilt auch dann noch, wenn eine Änderung des magnetischen Flusses bei statischen Magnetfeldern durch eine Bewegung, z. B. eine Verformung oder Drehung der Induktionsschleife, hervorgerufen wird (Abb. 2.2.4-1).

Der zweite Effekt ist beispielsweise bei einer Kondensatoraufladung (Abb. 2.2.4-2) wirksam. Das zeitlich veränderliche Feld wirkt wie eine Fortsetzung des Stroms in den Drähten zu den beiden Kondensatorplatten. Der Kondensatorbereich ist genau so wie der stromdurchflossene Draht von geschlossenen magnetischen Feldlinien umgeben.

2.2.4.2. Elektromagnetische Wellen. Zu den wichtigsten zeitabhängigen elektromagnetischen Feldern zählen die elektromagnetischen Wellen

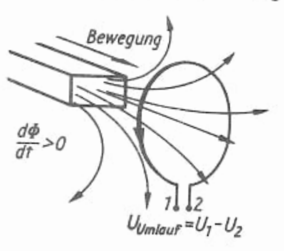


Abb. 2.2.4-1 Induktion bei feststehender Leiterschleife

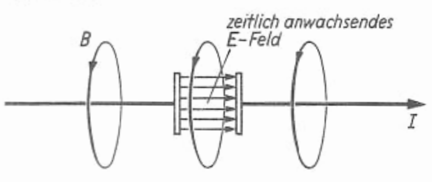


Abb. 2.2.4-2 Erzeugung eines Magnetfelds durch ein zeitlich veränderliches elektrisches Feld



Abb. 2.2.4-3 Linear polarisierte elektromagnetische Welle mit definierter Wellenlänge

(↑ 2.3.). Diese können sich ohne Vermittlung von elektrischen Ladungen oder Strömen im Vakuum ausbreiten. Die von MAXWELL in Weiterführung der Gedankengänge von FARADAY entwickelte Theorie elektromagnetischer Felder sagte die Existenz der wenig später von H. HERTZ experimentell nachgewiesenen elektromagnetischen Wellen voraus. Aus der Theorie ergibt sich ihre *Ausbreitungsgeschwindigkeit* zu $c = 1/\sqrt{\varepsilon_0\mu_0} = 2{,}9979 \cdot 10^8\ \mathrm{ms^{-1}}$. Die beiden Konstanten ε_0 und μ_0 (↑ 2.2.3.1. und 2.2.3.4.) können durch Experimente mit statischen Feldern bestimmt werden. Der so erhaltene Wert für c stimmt genau überein mit der durch direkte Messung bestimmbaren *Lichtgeschwindigkeit* im Vakuum, die sich hier als Ausbreitungsgeschwindigkeit von elektromagnetischen Wellen mit beliebiger Wellenlänge (↑ 2.2.4.3.) erweist. Damit ist die *elektromagnetische Lichttheorie* begründet. Die Tatsache, daß die Ausbreitungsgeschwindigkeit unabhängig von der Wellenlänge ist, stellt eine wichtige Besonderheit elektromagnetischer Wellen im Vakuum dar.

Am Zustandekommen einer elektromagnetischen Welle sind zeitabhängige elektrische und magnetische Felder gleichberechtigt beteiligt. Sie erzeugen sich aufgrund ihrer Kopplungen (↑ 2.2.4.1.) gewissermaßen gegenseitig und ermöglichen so die Ausbreitung der Welle. Bei einer ebenen Welle stehen sowohl der elektrische als auch der magnetische Feldvektor senkrecht auf der Ausbreitungsrichtung und senkrecht zueinander. Die elektromagnetischen Wellen sind *transversale Wellen*. Der elektrische und magnetische Anteil ihres Energieinhalts sind gleich groß. Es gilt $E = cB$. Wegen des allgemeinen Zusammenhangs zwischen elektrischem und magnetischem Feld genügt es, das Verhalten eines von beiden zu beschreiben.

Bleiben die Richtung des elektrischen Felds und damit auch die des magnetischen Felds konstant, so heißt die Welle *linear polarisiert* (Abb. 2.2.4-3). Durch Überlagerung zweier linear polarisierter Wellen entsteht im allgemeinen eine *elliptisch polarisierte* Welle (s. a. 2.1.4.3.), bei der an einem festen Ort die Spitzen des elektrischen und des magnetischen Feldvektors auf zwei ähnlichen, um 90° gegeneinander gedrehten Ellipsen umlaufen. Ein wichtiger Spezialfall hiervon ist die *zirkulare Polarisation*, bei der die Spitzen der Feldvektoren auf Kreisen laufen. Blickt man in Ausbreitungsrichtung und erfolgt dann die Drehung der Feldvektoren nach rechts, so heißt die Welle rechtszirkular polarisiert. Zirkular polarisiertes Licht entsteht bei Überlagerung von zwei in zueinander senkrechten Ebenen linear polarisierten Wellen gleicher Amplitude, die um eine halbe Wellenlänge gegeneinander verschoben sind.

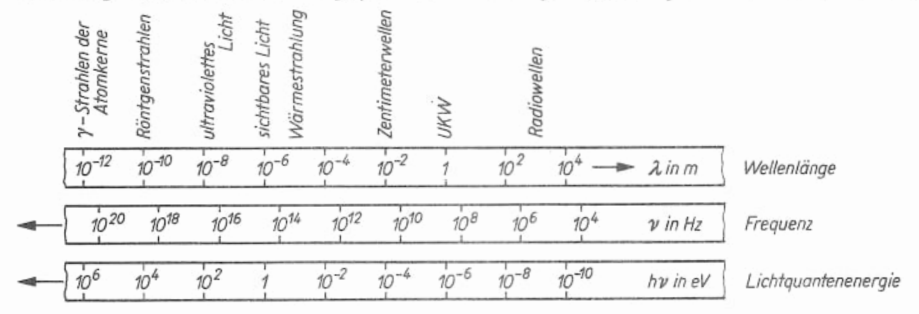


Abb. 2.2.4-4 Größenordnungen im elektromagnetischen Spektrum

2.2.4.3. Elektromagnetisches Spektrum. Das sichtbare Licht stellt nur einen ganz kleinen Ausschnitt aus der Gesamtheit aller elektromagnetischen Wellen dar (Abb. 2.2.4-4). Nach längeren Wellen schließt sich der infrarote, nach kürzeren der ultraviolette Spektralbereich an. Bis in diese Bereiche konnten noch die aus der Optik des sichtbaren Lichts bekannten experimentellen Methoden ausgedehnt werden, insbesondere beweisen direkte Interferenz- und Beugungsexperimente (↑ 2.3.2.) die Wellennatur der Strahlung. Für Wellen im Mikrowellenbereich, also Dezimeter- bis Millimeterwellen, stehen sehr genaue elektronische Meßmethoden, u. a. die gesamte Radartechnik zur Verfügung. Für Röntgenstrahlen dienen die Kristallgitter als Beugungsgitter (↑ 6.4.2.3.1.).

Mit zunehmender Frequenz treten die Welleneigenschaften zurück und die Strahlung verhält sich mehr wie ein Korpuskelstrahl. Der Grund hierfür liegt darin, daß bei sehr kurzen Wellenlängen keine Beugungserscheinungen mehr beobachtet werden, weil entsprechend kleine Hindernisse nicht zur Verfügung stehen. Gleichzeitig nimmt die Energie $h\nu$ eines Lichtquants (↑ 1.2.2.), die bei kleinen Frequenzen meist nicht ausreicht, um beobachtbare Effekte her-

vorzurufen, mit der Frequenz zu, so daß bei kurzwelliger Strahlung schon ein einziges Lichtquant wesentliche Umwandlungen an atomaren Systemen auslöst.

2.2.4.4. Energieströmung und Impuls. Bei zeitabhängigen Feldern ändert sich auch die räumliche Verteilung der Feldenergie. Es findet also eine Umverteilung, eine Strömung der Feldenergie statt. Der *Energiestrom* durch eine gegebene Fläche wird definiert als Quotient $\Delta E_F/\Delta t$ aus der Feldenergie ΔE_F, die in der (kurzen) Zeit Δt durch diese Fläche hindurchströmt; die *Energiestromdichte* als Energiestrom pro Fläche. Diese Begriffe werden genauso wie die Begriffe elektrischer Strom bzw. elektrische Stromdichte (↑ 2.2.2.2.) gebildet. Der Vektor $\boldsymbol{S}$ der Energiestromdichte beschreibt Stärke und Richtung der Energieströmung. Beim elektromagnetischen Feld gilt $\boldsymbol{S} = \mu_0^{-1}\boldsymbol{E} \times \boldsymbol{B}$, man bezeichnet $\boldsymbol{S}$ als *Poyntingvektor*.

Bei einer *ebenen Welle* mit einheitlicher Ausbreitungsrichtung strömt die Feldenergie natürlich in Ausbreitungsrichtung. Tatsächlich zeigt der Poyntingvektor aufgrund der für ebene elektromagnetische Wellen typischen Richtungsbeziehungen in diese Richtung. Für seinen Betrag gilt dann $S = wc$ analog zur Darstellung $j = \varrho v$ der elektrischen Stromdichte (↑ 2.2.2.2.), wie es bei dem mit Lichtgeschwindigkeit erfolgenden Transport der mit der Energiedichte w verteilten Feldenergie sein muß.

Mit der Energieströmung ist ein *Impuls des Felds* verknüpft. Für eine Welle (Wellenpaket) mit einheitlicher Ausbreitungsrichtung gilt der Zusammenhang $E_F = cP_F$ zwischen der gesamten Energie E_F und dem gesamten Impuls P_F der Welle. Das ist genau der relativistische *Energie-Impuls-Zusammenhang* für ein Teilchen ohne Ruhmasse (↑ 2.1.3.5.). Dementsprechend sind die den elektromagnetischen Wellen zugeordneten Teilchen, die *Lichtquanten* (↑ 1.2.2.), Teilchen ohne Ruhmasse.

Bei der *Absorption* einer elektromagnetischen Welle wird ihre Energie und ihr Impuls auf das absorbierende System übertragen. Bei vollständiger *Reflexion* wird der doppelte Impuls übertragen, da der Impuls der Welle durch die Reflexion sein Vorzeichen wechselt. Die Impulsübertragung führt genau wie bei der Reflexion von Gasatomen an der Wand des Gefäßes zu einem Druck auf die absorbierende bzw. reflektierende Wand. Dieser als *Strahlungs- oder Lichtdruck* bezeichnete Druck konnte erstmals von LEBEDEV experimentell nachgewiesen werden.

2.2.5. Allgemeines Energieprinzip

2.2.5.1. Energiesatz. Jeder physikalischen Erscheinung kann eine skalare Größe Energie so zugeordnet werden, daß die Summe aller Energien sich zeitlich nicht ändert, obwohl die beteiligten Teilchen und Felder sich bewegen und verändern. Diese Aussage, die am Beispiel des Systems „geladene Teilchen und elektromagnetisches Feld" schon erläutert wurde (↑ 2.2.3.7.), ist der Inhalt des allgemeinen Erhaltungssatzes der Energie, kurz Energiesatz genannt. Dieser Satz gilt nur, wenn die Energien aller beteiligten Teilchen, Felder usw. berücksichtigt werden. Scheinbare Verletzungen des Energiesatzes konnten bisher immer dadurch aufgeklärt werden, daß die Energie eines zunächst vergessenen Teilchens oder Felds in die Energiebilanz aufgenommen wurde. Ein historisches Beispiel dafür ist die Einführung des Neutrinos in die Theorie des β-Zerfalls von Atomkernen (↑ 3.3.1.2.). Die Existenz ungeladener und daher direkt schwer nachweisbarer Elementarteilchen wird meist nur aus der Gültigkeit von Energie- und Impulssatz gefolgert.

Die enorme *praktische Bedeutung des Energiebegriffs* – man denke an Energiehaushalt, Energie und technische Entwicklung – beruht nicht allein auf der Eigenschaft der Energie, eine Erhaltungsgröße zu sein. Dafür ist in gleichem Maße die *Additivität der Energie* wesentlich, durch die die Energie erst zum universellen *Maß für die Quantität* einer beliebigen Erscheinung wird. Additivität heißt, daß die kinetische Energie zweier Massen gleich der Summe der einzelnen kinetischen Energien ist, daß die Energie zweier räumlich getrennter elektrischer Felder gleich der Summe der einzelnen Feldenergien ist.

2.2.5.2. Energieformen. Um für einen physikalischen Vorgang die Energiebilanz aufstellen zu können, muß man die Energien aller beteiligten Objekte kennen. Wieviel wesentlich verschiedene Energiearten man einführen muß, hängt zum Teil vom Stand der Naturerkenntnis ab. Bleibt man bei den äußeren Erscheinungsformen stehen, so treten folgende Energien auf: kinetische, potentielle, elastische, elektrische und magnetische Energie, Gravitationsenergie, Lichtenergie, Wärmeenergie, chemische Energie, Kernenergie. Die moderne Physik, insbesondere die Atomphysik, ermöglicht eine wesentliche Vereinfachung dieses Bilds. Die Zurückführung der potentiellen Energie auf andere Energieformen wurde bereits besprochen (↑ 2.1.5.4.). Licht ist ein spezieller Fall elektromagnetischer Wellen (↑ 2.2.4.2.); die *Lichtenergie* ist damit auf die elektrische und magnetische Energie dieser Wellen zurückgeführt. *Wärmeenergie* ist identisch mit der inneren kinetischen und potentiellen Energie eines aus vielen Teilchen aufgebauten makroskopischen Systems.

Mit den Hilfsmitteln der Quantentheorie wird

die *chemische Energie*, d. h. die Energie der chemischen Bindungen (↑ 5.1.1.), als Änderung der kinetischen und der elektrostatischen Wechselwirkungsenergie der Elektronen im Molekül gegenüber ihrer Energie in den getrennten Atomen erkannt. Die Energie einer *elastischen Deformation* besteht aus der gesamten Änderung der potentiellen Energie, die den einzelnen Atomen des Materials aufgrund ihrer gegenseitigen Wechselwirkung zukommt. Diese Wechselwirkungsenergie ist ähnlich wie die Energie der chemischen Bindung auf Änderungen der Elektronenenergien durch die gegenseitige Störung der Atome zurückführbar. Die bei der *Kernspaltung* frei werdende Energie ist zum größten Teil elektrostatische Energie. Diese ist größer, wenn alle Protonen im Ausgangskern vereinigt sind, als wenn sie auf zwei getrennte Teilkerne verteilt sind. Beim Auseinanderfliegen werden die Bruchstücke durch die zwischen ihnen wirkende elektrische Abstoßungskraft beschleunigt. Dabei wandelt sich elektrostatische Energie in kinetische Energie um.

Eine besondere Rolle spielt die *Ruhenergie* (↑ 2.1.3.6.). Während sie bei zusammengesetzten Teilchen bzw. makroskopischen Objekten als Maß des Energieinhalts des Systems zumindest teilweise auf andere Energieformen zurückführbar ist, wird diese Zurückführung bei Elementarteilchen sehr problematisch, so daß die Ruhenergie als selbständige Energieform erscheint. Eine scharfe Grenze zwischen Elementarteilchen und zusammengesetzten Teilchen existiert jedoch nicht.

Diese Beispiele zeigen: Bei physikalischen Vorgängen werden verschiedene Energiearten ineinander umgewandelt. Als wesentlich verschiedene Energiearten bleiben übrig: die Ruhenergie und die kinetische Energie von bewegten Teilchen, die Feldenergie elektromagnetischer Felder, die Gravitationsenergie der Schwerefelder, die Energien von Kernkraftfeldern. Für das Verständnis der Atomstruktur und des atomaren Aufbaus der Materie spielen fast nur kinetische Energien und die Energie von elektrischen Feldern eine Rolle.

2.2.5.3. Energieeinheiten. In verschiedenen Gebieten der Physik bzw. bei Betrachtung verschiedener Energieformen sind verschiedene Energieeinheiten gebräuchlich. Die gesetzliche Einheit 1 Joule = 1 J = 1 Nm = 1 kgm^{2}s^{-2} ist an dem mechanischen Arbeits- und Energiebegriff orientiert, entsprechend der Arbeitsdefinition (↑ 2.1.5.4.) ist sie als Produkt der Krafteinheit 1 N mit der Längeneinheit 1 m definiert. Die Leistungseinheit 1 Watt = 1 W ist durch 1 J = 1 Ws definiert. Andere Energieeinheiten werden durch Definition an die gesetzliche Einheit Joule angeschlossen. Aufgrund der Definition der Einheit 1 Volt der elektrischen Spannung (↑ 2.2.3.2.) gilt 1 W = 1 VA. Wenn ein Strom von 1 A eine Strecke durchfließt, zwischen deren beiden Enden eine Spannung von 1 V liegt, so nehmen die Ladungsträger innerhalb einer Sekunde eine Energie von 1 J = 1 Ws auf. Durchläuft ein Teilchen mit einer Elementarladung eine Spannungsdifferenz von 1 V, so wird die kinetische Energie 1 Elektronenvolt = 1 eV = $1{,}602 \cdot 10^{-19}$ J.

2.3. Wellen

2.3.1. Kenngrößen und Eigenschaften

2.3.1.1. Begriffsbildung. Die Ausbreitung von Wellen wird bei physikalisch ganz verschiedenen Systemen beobachtet, z. B. Wasserwellen, Schallwellen, Seilwellen, Lichtwellen, Radiowellen und Materiewellen. Alle diese Ausbreitungsphänomene zeigen gemeinsame Züge. Der Ausbreitungsvorgang kann beschrieben werden, indem man den zeitlichen Ablauf eines räumlichen Felds (↑ 2.2.1.) angibt. Bei Wasser- und Schallwellen handelt es sich um Strömungs- oder Geschwindigkeitsfelder und um Druckfelder, bei Licht- und Radiowellen um elektromagnetische Felder. Beschränkt man den Begriff Welle nicht unbedingt auf einen zeitlich und räumlich periodischen Ablauf, so kann eine scharfe Grenze zwischen einer Welle und einem zeitabhängigen Feld nicht gezogen werden.

Das Quadrat der jeweiligen Feldgröße bestimmt die *Intensität* der Welle. Von der Intensität der Lichtwelle hängt z. B. die Schwärzung eines Films bei Bestrahlung ab.

Der räumliche Feldverlauf kann für verschiedene Wellen ganz verschieden sein. Stoßwellen haben eine scharfe Wellenfront. Kugelwellen laufen von einem Zentrum fort. Der Feldverlauf zu einer festen Zeit, eine Momentaufnahme der Welle, kann stets dadurch veranschaulicht werden, daß man die Flächen konstanten Wellenfelds zeichnet. Die Wellenbewegung ist dann mit einer Bewegung, im allgemeinen auch mit einer Deformation dieser Flächen verknüpft.

2.3.1.2. Ebene Wellen. Die einfachsten Wellen sind die ebenen Wellen. Bei ihnen sind die Flächen konstanten Wellenfelds parallele Ebenen, die senkrecht auf der Ausbreitungsrichtung der Wellen stehen und sich in dieser Richtung verschieben. Zu jedem Zeitpunkt, also für jede „Momentaufnahme" der Welle, hängt also der Feldverlauf nur von einer einzigen räumlichen Koordinate, nämlich der in Ausbreitungsrichtung ab. Orientiert man die x-Achse in Ausbreitungsrichtung, so hängt die Feldgröße $f(x, t)$ nicht von y und z ab. Beim Vergleich zweier Zeitpunkte t_1 und t_2 (Abb. 2.3.1-1) ergeben sich verschiedene „Momentaufnahmen".

Beim Vergleich zweier Momentaufnahmen kann die Form des Feldverlaufs, die „Signalform" sich als verschieden herausstellen. Das Signal wird dann bei der Ausbreitung „verzerrt". Nur wenn eine einheitliche Ausbreitungsgeschwindigkeit existiert, bleibt die Signalform erhalten. Dieser für die Anwendung bedeutsame Fall tritt bei den elektromagnetischen Wellen im Vakuum auf, während z. B. Materiewellen eine wesentliche Verzerrung erfahren.

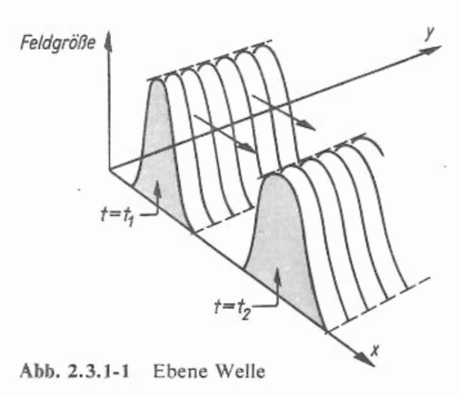


Abb. 2.3.1-1 Ebene Welle

2.3.1.3. Frequenz und Wellenlänge. Besonders einfach und wichtig sind solche ebenen Wellen, die zeitlich und räumlich harmonisch sind. Bei einer Momentaufnahme ändert sich dann der Feldverlauf beim Weiterschreiten in Ausbreitungsrichtung nach einer Sinus- oder Kosinuskurve. Der Abstand zweier Punkte, zwischen denen eine volle räumliche Periode liegt, heißt *Wellenlänge* λ. Die Wellenausbreitung besteht darin, daß sich der gesamte Feldverlauf im Laufe der Zeit mit konstanter Geschwindigkeit in Ausbreitungsrichtung bewegt. An einem festen Punkt zieht dann Wellenzug auf Wellenzug vorüber, und man beobachtet einen in der Zeit periodischen Verlauf, eine harmonische Schwingung (↑ Abb. 2.1.4-3). Die Zeitdauer einer vollen Schwingungsperiode heißt *Schwingungsdauer* T, das Reziproke davon ist die *Frequenz* $\nu = 1/T$ (↑ 2.1.4.3.). Während einer Schwingungsdauer T zieht gerade eine volle Wellenlänge an dem festen Punkt vorüber. Die Geschwindigkeit, mit der der gesamte Wellenzug sich nach rechts verschiebt, ist die *Phasengeschwindigkeit* $v_{ph} = \lambda/T = \lambda\nu$ der Welle. Für die Schwingungsdauer erhält man $T = \nu^{-1} = \lambda/v_{ph}$.

Eine sich in x-Richtung ausbreitende harmonische ebene Welle wird mathematisch durch den Verlauf der Feldgröße $= A \cos\left[2\pi\left(\frac{x}{\lambda} - \frac{t}{T}\right)\right]$ $= A \cos(kx - \omega t)$ beschrieben. Der Maximalwert A heißt *Amplitude*, das Argument der Kosinusfunktion ist die vom Ort x und von der Zeit t abhängige *Phase* der Welle. Die Phase ist in der zweiten Form mittels der *Kreisfrequenz* $\omega = 2\pi\nu$ (↑ 2.1.4.3) und der *Wellenzahl* $k = 2\pi/\lambda$ dargestellt. Der Zusammenhang $\lambda\nu = v_{ph}$ zwischen Wellenlänge und Frequenz kann in der Form $\omega = v_{ph}k$ als Relation zwischen ω und k dargestellt werden.

Die Wellenlängen sichtbaren Lichts im Vakuum liegen etwa zwischen $0{,}4 \cdot 10^{-6}$ und $0{,}8 \cdot 10^{-6}$ m. Da 10^{-6} m $= 0{,}001$ mm, sind sie verglichen mit den im täglichen Leben bekannten Längen recht klein. Die zugehörigen Frequenzen liegen zwischen $0{,}4 \cdot 10^{15}$ Hz und $0{,}7 \cdot 10^{15}$ Hz. Die Lichtschwingungen erfolgen danach unvorstellbar rasch. Dagegen liegen die Frequenzen des Hörschalls wesentlich niedriger, nämlich im Bereich von etwa 20 Hz bis 20 kHz, und die zugehörigen Wellenlängen (in Luft) liegen im Bereich von 10 m bis 1 cm.

Eine Welle, die sich in einer beliebigen, durch einen Einheitsvektor $\boldsymbol{e}$ gekennzeichneten Richtung ausbreitet, wird durch den mathematischen Ausdruck: Feldgröße $= A \cos(k\boldsymbol{er} - \omega t)$ $= A \cos(\boldsymbol{kr} - \omega t)$ beschrieben. Das folgt daraus, daß $\boldsymbol{er}$ gerade die Komponente des Ortsvektors $\boldsymbol{r}$ in Ausbreitungsrichtung darstellt. Der *Wellenzahlvektor* $\boldsymbol{k} = k\boldsymbol{e}$ faßt die Angaben über Richtung und Wellenlänge $\lambda = 2\pi/k$ zusammen.

In einem die Energie des Wellenfelds absorbierenden Medium kann sich nur eine gedämpfte Welle ausbreiten. Prinzipiell existieren dabei zwei Möglichkeiten:

Das Feld bleibt zeitlich rein harmonisch, erhält aber eine räumliche Dämpfung derart, daß die Amplitude in Ausbreitungsrichtung abnimmt. Da die Frequenz im allgemeinen vorgegeben ist und die Welle von einer Grenzfläche her in das dämpfende Medium eindringt, ist dieser Fall typisch.

Das Feld erhält eine zeitliche Dämpfung derart, daß sich die Amplitude (im ganzen Raum gleichmäßig) im Zeitablauf verringert. Dieser Fall ist bei echten Ausbreitungsvorgängen nicht realisierbar, aber typisch für gedämpfte Eigenschwingungen (↑ 2.3.3.).

2.3.1.4. Dispersion und Brechungsindex. Im allgemeinen hängt die Phasengeschwindigkeit einer harmonischen Welle von deren Wellenlänge ab, z. B. bei Licht in Substanzen oder bei Materiewellen. Der Zusammenhang zwischen Frequenz ν und reziproker Wellenlänge λ^{-1} bzw. zwischen Kreisfrequenz ω und Wellenzahl k ist nur bei konstanter Phasengeschwindigkeit linear, im allgemeinen ist er weitgehend beliebig. Der für die jeweilige Wellenausbreitung charakteristische Zusammenhang zwischen λ und ν wird als *Dispersionsbeziehung* bezeichnet und meist als Funktion $\omega = \omega(k)$ dargestellt. Eine lineare Dispersionsbeziehung $\omega \sim k$

bedeutet eine konstante Phasengeschwindigkeit. Nur in diesem Fall ist mit der Ausbreitung keine Signalverzerrung verbunden.
Bei gegebener Frequenz, die sich beim Übergang der Welle in ein anderes Ausbreitungsmedium nicht ändert, hängt die Wellenlänge vom jeweiligen Medium ab. Als *Brechungsindex* n des Mediums bezeichnet man eine dimensionslose positive Größe, deren Verhältnis $n_1/n_2 = \lambda_2/\lambda_1 = v_{ph}^{(2)}/v_{ph}^{(1)}$ beim Vergleich zweier Medien das Verhältnis der Wellenlängen bzw. Phasengeschwindigkeiten bestimmt. Großer Brechungsindex bedeutet kleine Wellenlänge bzw. kleine Phasengeschwindigkeit. Da nur Verhältnisse von Brechungsindizes physikalische Bedeutung haben, ist der Absolutwert von n im allgemeinen belanglos.
Falls es ein ausgezeichnetes Ausbreitungsmedium gibt, wie das Vakuum für elektromagnetische Wellen, setzt man den zugehörigen Brechungsindex gleich 1. Der Brechungsindex für eine Lichtwelle ist also durch $n = c/v_{ph}$ definiert. Er enthält dann die gleiche Information wie die Dispersionsrelation $\omega(k)$, denn es gilt $\omega = v_{ph}k = ck/n$.
Im allgemeinen hängt der Brechungsindex von der Frequenz der Welle ab. Diese Erscheinung heißt *Dispersion.*

2.3.1.5. Gruppengeschwindigkeit. Eine harmonische ebene Welle ist räumlich streng periodisch mit der Wellenlänge λ und daher nicht auf einen bestimmten Raumbereich lokalisiert. Sie ist zur Signalübertragung ungeeignet. Um Signalstrukturen mit direkt meßbarer Ausbreitungsgeschwindigkeit zu erhalten, muß die Welle „moduliert" werden. Dies geschieht durch Überlagerung von Wellen unterschiedlicher Wellenlänge. Als einfaches Beispiel kann eine sogenannte *Schwebung* (Abb. 2.3.1-2) dienen,

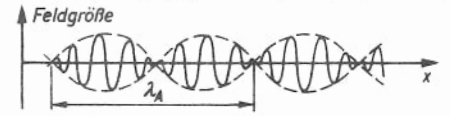


Abb. 2.3.1-2 Schwebung

die durch Überlagerung zweier harmonischer Wellen mit den Wellenlängen λ_1 und λ_2 und geringem Wellenlängenunterschied $\Delta\lambda$ entsteht. Die resultierende amplitudenmodulierte Welle zeigt eine rasche Oszillation mit der mittleren Wellenlänge $\bar{\lambda}$, wobei die Amplitude mit der Wellenlänge $\lambda_A = \lambda_1\lambda_2/\Delta\lambda$ moduliert ist. Die Ortsabhängigkeit der Amplitude ist also durch eine Wellenzahl $k_A = 2\pi/\lambda_A = 2\pi|\lambda_1^{-1} - \lambda_2^{-1}| = |k_1 - k_2|$ charakterisiert. Entsprechend oszilliert die Amplitude zeitlich mit der Differenzfrequenz $\nu_A = |\nu_1 - \nu_2|$. Daraus folgt, daß $v_A = \Delta\omega/\Delta k$ die Geschwindigkeit ist, mit der sich das Amplitudengebirge der Schwebung in Ausbreitungsrichtung verschiebt.
Durch Überlagerung von mehr als zwei Wellen können *kompliziertere Signalformen* erzeugt werden. Durch Superposition harmonischer Wellen mit kontinuierlich variierendem λ bzw. k können völlig beliebige *Wellenpakete* (↑ 2.3.1.6.) aufgebaut werden. Ist das bei der Überlagerung benutzte Wellenlängenintervall so schmal, daß die Dispersionsbeziehung in diesem Intervall durch einen linearen Verlauf ersetzt werden kann, so ist $v_{gr} = d\omega/dk$ die sogenannte *Gruppengeschwindigkeit*, mit der sich das Wellenpaket als Ganzes bewegt.
Falls $\omega = v_{ph}k$ mit *konstanter* Phasengeschwindigkeit v_{ph}, also eine lineare Dispersionsbeziehung, gilt, stimmen Phasen- und Gruppengeschwindigkeit überein. In allen anderen Fällen können sie wesentlich verschieden sein. Die Gruppengeschwindigkeit kann als Signalgeschwindigkeit höchstens gleich der Vakuumlichtgeschwindigkeit sein. Für elektromagnetische Wellen im Vakuum ist $v_{gr} = v_{ph} = c$.
Bei nichtlinearer Dispersionsbeziehung ist $d\omega/dk$ auch in einem kleinen Wellenlängenintervall nicht genau konstant. Die verschiedenen Anteile eines Wellenpakets laufen dann unterschiedlich schnell, was schließlich zum *Zerfließen des Wellenpakets*, d. h. zu seiner unbegrenzten Verbreiterung führt.

2.3.1.6. Spektraldarstellung von Wellenpaketen. Zu jedem beliebigen, zunächst festgehaltenem Zeitpunkt, also für eine „Momentaufnahme" der Welle, kann der Feldverlauf durch Überlagerung rein harmonischer sin kx- und cos kx-Verläufe dargestellt werden. Besitzt der Feldverlauf selbst keine räumliche Periodizität, so ist es notwendig, die Wellenzahl k kontinuierlich zu variieren; die Überlagerung hat dann die Gestalt eines *Fourierintegrals*, man vergleiche auch die analogen Betrachtungen zur Darstellung einer gedämpften Schwingung (↑ 2.1.6.1.). Die mathematische Darstellung ist besonders übersichtlich mit der komplexen Zusammenfassung $\exp(ikx) = \cos kx + i \sin kx$. Dann muß man gleichzeitig positive und negative k-Werte benutzen, um entsprechend $\cos kx = \frac{1}{2}(e^{ikx} + e^{-ikx})$ und $\sin kx = \frac{1}{2i}(e^{ikx} - e^{-ikx})$ die reellen Kosinus- und Sinusfunktionen wiedergeben zu können. Dann gilt

$$f(x) = \int_{-\infty}^{+\infty} \frac{dk}{2\pi} f(k)\, e^{ikx}.$$

Reelle Funktionen führen auf eine symmetrische Funktion $f(k) = f(-k)$. Die im allgemeinen komplexe Funktion $f(k)$ heißt die *Fouriertransformierte* von $f(x)$; sie kann entsprechend $f(k) = \int_{-\infty}^{+\infty} dx\, f(x)\, e^{-ikx}$ aus $f(x)$ eindeutig be-

stimmt werden. Durch die Phase der komplexen Größe $f(k)$ wird die Interferenz (↑ 2.3.2.) der überlagerten „Wellen", d. h. der zunächst nur x-abhängigen $\exp(ikx)$-Verläufe, wesentlich beeinflußt.

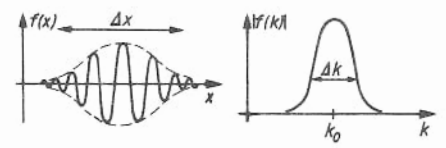


Abb. 2.3.1-3 Breite eines Wellenpakets und seiner Fouriertransformierten

Um ein Wellenpaket der räumlichen Breite Δx herzustellen, braucht man ebene Wellen, deren Wellenzahlen mindestens ein Intervall Δk überdecken (Abb. 2.3.1-3). Wenn $\Delta x\, \Delta k \approx 1$ ist, entstehen zwischen den überlagerten „Wellen" auf der Strecke Δx Phasenunterschiede von der Größenordnung 2π bzw. π, d. h., es treten Gangunterschiede von einer bzw. einer halben Wellenlänge auf. Daher können sich die zum Wellenpaket überlagerten Wellen außerhalb Δx durch Interferenz auslöschen. Ein scharf lokalisiertes Wellenpaket enthält ebene Wellen mit sehr verschiedenen Wellenlängen. Ein sehr ausgedehntes Wellenpaket braucht nur Wellenlängen aus einem kleinen Intervall zu enthalten. Es kann häufig durch eine ebene Welle (nicht lokalisiert, $\Delta\lambda = 0$) idealisiert werden.

Der *Zeitablauf des Wellenpakets* entsteht dadurch, daß jede einzelne harmonische Komponente $\exp(ikx)$ *ihren* zeitlich periodischen Faktor $\exp(-i\omega(k)\,t)$ erhält, wobei die Frequenz $\omega = \omega(k)$ nach der Dispersionsbeziehung (↑ 2.3.1.4.) eindeutig zugeordnet ist. Die Funktion $\exp(ikx - i\omega t)$ stellt eine komplexe ebene Welle vollständig dar. Ihren rein ortsabhängigen Anteil $\exp(ikx)$ bezeichnet man jedoch ebenfalls als ebene Welle, wobei man sich den Zeitfaktor $\exp(-i\omega t)$ gewissermaßen dazudenken muß. Der Realteil von $\exp[i(kx - \omega t)]$ stellt ebenso wie der Imaginärteil eine reelle harmonische Welle (↑ 2.3.1.2.) dar. Das zeitabhängige Wellenpaket $f(x, t) = \int \frac{dk}{2\pi} f(k)\, e^{i(kx-\omega t)}$ verhält sich so wie diskutiert (↑ 2.3.1.5.).

Definiert man als *Durchlaufzeit* Δt des Wellenpakets die Zeit, in der es einen festen Punkt passiert, so muß $\Delta t = \Delta x / v_{gr}$ gelten. Mit $v_{gr} = d\omega/dk \approx \Delta\omega/\Delta k$ erhält Δt die Form $\Delta t = \frac{\Delta x\, \Delta k}{\Delta\omega} \approx \frac{1}{\Delta\omega}$ bzw. $\Delta t\, \Delta\omega \approx 1$, wenn man $\Delta x\, \Delta k \approx 1$ beachtet. Zwischen Frequenzunschärfe und „zeitlicher Länge" des Pakets besteht also eine gleichartige Reziprozitätsbeziehung wie zwischen Δk und der räumlichen Ausdehnung Δx.

2.3.1.7. Reflexion und Brechung. An einer Grenzfläche zwischen zwei Medien mit unterschiedlichen Ausbreitungseigenschaften wird eine Welle teilweise reflektiert und teilweise durchgelassen. Die Intensität (die Energiestromdichte) der Primärwelle teilt sich auf in die Intensität der reflektierten und die der durchgelassenen Welle. Fällt eine (beliebige) ebene Welle auf eine ebene Grenzfläche (Abb. 2.3.1-4), so ist auch die reflektierte Welle eben. Die Ausbreitungsrichtungen beider Wellen und die Grenzflächennormale liegen in einer Ebene. Einfallswinkel α_1 und Ausfallswinkel α_1' sind gleich. Falls für alle in der Primärwelle vertretenen Wellenlängen λ_1 (bzw. für alle Frequenzen) das Verhältnis der Wellenlängen λ_1/λ_2 in beiden Medien konstant ist, ist auch die durchgelassene Welle eben. Ihre Ausbreitungsrichtung liegt dann ebenfalls in der erwähnten Ebene, bildet aber einen anderen Winkel α_2 mit der Grenzflächennormale. Diese Richtungsänderung wird als Brechung bezeichnet. Es gilt das *Brechungsgesetz*

$$(1)\quad \frac{\sin\alpha_1}{\sin\alpha_2} = \frac{\lambda_1}{\lambda_2} = \frac{v_{ph}^{(1)}}{v_{ph}^{(2)}} = \frac{n_2}{n_1},$$

wobei n_1 und n_2 die Brechungsindizes (↑ 2.3.1.4.) beider Medien sind. Bei Lichtbrechung heißt das Medium mit dem größeren Brechungsindex optisch dichter, das mit dem kleineren optisch dünner.

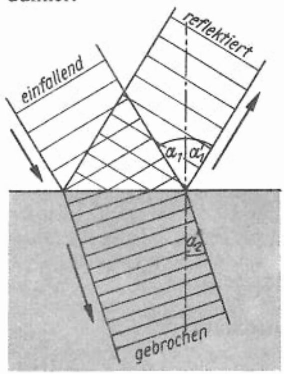


Abb. 2.3.1-4 Reflexion und Brechung einer ebenen Welle

Hängt das Verhältnis n_2/n_1 von der Primärwellenlänge λ_1 ab, so werden harmonische Wellen mit verschiedener Wellenlänge unterschiedlich gebrochen. Darauf beruht die *spektrale Zerlegung* des Lichts durch ein Prisma. Die Frequenz bleibt beim Übergang ins andere Medium stets unverändert.

Die allgemeinen Richtungsaussagen sowie das Brechungsgesetz in der Form, daß das Verhält-

nis $\sin\alpha_1/\sin\alpha_2$ eine vom Einfallswinkel α_1 unabhängige Größe darstellt, gelten auch für *Teilchenstrahlen*, wenn Teilchen mit betragsmäßig einheitlicher Geschwindigkeit aus einem Raumbereich mit dem konstanten Potential V_1 (↑ 2.1.5.6.) in einen anderen mit dem Potential V_2 übergehen. Die Tangentialkomponente der Teilchengeschwindigkeit bleibt unverändert, während sich die Normalkomponente so verändert, daß dem mechanischen Energiesatz entsprechend auf beiden Seiten der Grenzfläche die Summe aus kinetischer und potentieller Energie denselben Wert hat. Daher ändert sich die Strahlrichtung. Es gilt $\sin\alpha_1/\sin\alpha_2 = v_2/v_1$, wenn v_1 und v_2 die Teilchengeschwindigkeiten in beiden Bereichen sind. Ein wichtiger Unterschied zwischen dem Verhalten von Teilchen und Wellen bei Reflexion und Brechung besteht darin, daß Teilchen *entweder* (vollständig) *reflektiert oder* (nur) *gebrochen* werden, je nachdem, ob ein Übergang ins andere Medium energetisch möglich ist oder nicht. Dagegen wird ein Teil der Welle immer reflektiert, auch wenn sich diese im zweiten Medium ausbreiten, also gewissermaßen bewegen kann.

2.3.1.8. Totalreflexion. Manche Medien erlauben in bestimmten Frequenzbereichen nicht die Ausbreitung einer Welle, bei der wirklich Energie durch das Medium hindurchströmt. Dieser Effekt ist nicht damit zu verwechseln, daß ein absorbierendes Medium die Welle dämpfen kann; er tritt am deutlichsten bei nichtabsorbierenden Medien auf. Fällt eine Primärwelle, deren Frequenz in einem solchen „verbotenen" Frequenzbereich liegt, auf die Grenzfläche des Mediums, so wird sie vollständig reflektiert, unabhängig vom Einfallswinkel. Diese Erscheinung heißt *Totalreflexion*, das Medium wirkt als ein idealer Reflektor.

Dabei kann das totalreflektierende Medium nicht völlig feldfrei bleiben, da dann das Wellenfeld beim Durchgang durch die Grenzfläche plötzlich (unstetig) auf Null abnehmen müßte. Daher bildet sich im reflektierenden Medium ein von der Grenzfläche aus nach innen exponentiell abklingender Feldverlauf aus, den man kaum noch als Welle bezeichnen kann und mit dem auch kein Energietransport in das Medium hinein verbunden ist. Bei schrägem Einfall erzwingen die Stetigkeitsbedingungen, daß das Feld längs der Grenzfläche auch im Medium genau so oszilliert wie außerhalb; senkrecht zur Grenzfläche verstärkt sich dabei der exponentielle Abfall. In jedem Falle kann die Welle eine hinreichend dünne Schicht „verbotenen Materials" durchdringen und „auf der anderen Seite" als durchgelassene Welle in Erscheinung treten. Die Totalreflexion an der Schichtoberfläche ist dann unvollständig.

Die gleiche Situation wie bei einem die Wellenausbreitung überhaupt nicht ermöglichenden Medium kann bei einem an sich „normalen" Medium durch hinreichend schrägen Einfall entstehen. Durch die erzwungene (räumliche) Oszillation des Felds parallel zur Grenzfläche (s. o.) bleibt für eine Oszillation und damit für die Ausbreitung senkrecht zur Grenzfläche gewissermaßen nichts übrig, so daß sich wiederum ein exponentiell abfallender Feldverlauf einstellt. Das Brechungsgesetz 2.3.(1) läßt erkennen, wann dieser Fall eintritt. Für $\lambda_2 > \lambda_1$ erreicht $\sin\alpha_2$ seinen Maximalwert $\pi/2$, d. h. die gebrochene Welle breitet sich parallel zur Grenzfläche aus, wenn α_1 den durch $\sin\alpha_T = \lambda_2/\lambda_1 = n_1/n_2$ bestimmten Winkel $\alpha_T < \pi/2$ erreicht. Vergrößert man α_1 über den *Grenzwinkel* α_T *der Totalreflexion* hinaus, so wird die gesamte einfallende Intensität reflektiert. Dieses zunächst etwas merkwürdige Verhalten der Welle ist analog zum Verhalten von Teilchen, die schräg auf eine Grenzfläche zu einem Raumbereich mit größerem Potential $V_2 > V_1$ treffen. Auch wenn ihre gesamte kinetische Energie größer ist als die Potentialdifferenz ΔV, können sie diese nur dann überwinden, wenn die von der Bewegung *normal* zur Grenzfläche herrührende kinetische Energie $mv_\perp^2/2$ größer als ΔV ist.

2.3.1.9. Dopplereffekt. Als Dopplereffekt wird die durch eine Relativbewegung von Quelle und Beobachter verursachte Frequenzänderung bezeichnet. Bewegen sich Quelle und Beobachter aufeinander zu, so wird eine Frequenzerhöhung, anderenfalls eine Erniedrigung beobachtet. Dieser Effekt kann zur Messung von Geschwindigkeiten, u. a. von künstlichen Satelliten, benutzt werden. Der Dopplereffekt von Schallwellen kann an der Tonhöhe der von schnell fahrenden Fahrzeugen erzeugten Geräusche, z. B. am Pfeifen einer Lokomotive, leicht wahrgenommen werden.

Existiert in einem gegebenen Bezugssystem eine Welle mit der Frequenz ν und der Wellenlänge λ, Phasengeschwindigkeit $v_{ph} = \nu\lambda$, so wird ein mit einer nichtrelativistischen Geschwindigkeit $v_B \ll c$ gegenüber diesem System bewegter Beobachter die gleiche Wellenlänge $\lambda' = \lambda$, aber eine veränderte Ausbreitungsgeschwindigkeit $v'_{ph} = v_{ph} \pm v_B$ registrieren; je nachdem, ob er sich auf die Quelle zu (+) oder von ihr weg (−) bewegt. Es folgt $\nu'/\nu = v'_{ph}/v_{ph} = 1 \pm v_B/v_{ph}$ für die vom bewegten Beobachter gemessene Frequenz ν'. Da er entweder „mit der Welle" oder „der Welle entgegen" läuft, passieren ihn in der Zeiteinheit weniger oder mehr Wellenzüge als einen ruhenden Beobachter. Im Extremfall kann er (falls $v_{ph} < c$ ist) genau so schnell laufen wie die Welle, so daß diese für ihn „stillsteht"; die Frequenz ν' wäre dann Null.

Bei großer Geschwindigkeit v_B werden relativistische Effekte wichtig. Das *exakte Transforma-*

tionsgesetz für die Frequenz lautet

$$(2)\qquad \frac{\nu'}{\nu} = \frac{1 \pm v_B/v_{ph}}{\sqrt{1 - v_B^2/c^2}} .$$

Für $v_B \ll c$ geht es in das einfachere nichtrelativistische Gesetz ohne den Wurzelfaktor über. Für *elektromagnetische Wellen* ist $v_{ph} = c$, so daß sich bei Beachtung der Regel $1 - \beta^2 = (1 - \beta)(1 + \beta)$ der Ausdruck

$$(3)\qquad \frac{\nu'}{\nu} = \sqrt{\frac{1 \pm v_B/c}{1 \mp v_B/c}}$$

ergibt. Im Gegensatz zur nichtrelativistischen Näherung liefert die exakte Theorie auch einen *transversalen Dopplereffekt* bei Bewegung des Beobachters senkrecht zur Ausbreitungsrichtung der Welle.

Existiert wie bei *elektromagnetischen Wellen im Vakuum* kein ausgezeichnetes Ausbreitungsmedium, so kann das Ausgangsbezugssystem stets das der ruhenden Quelle sein. Der Dopplereffekt hängt dann in Einklang mit dem Relativitätsprinzip (↑ 2.1.2.4.) nur von der Relativgeschwindigkeit zwischen Quelle und Beobachter ab. Gibt es dagegen wie bei *Schallwellen* ein ausgezeichnetes Medium, so ist die entstehende Welle nur dann einfach, d. h. kugelsymmetrisch von der Quelle auslaufend, wenn die Quelle im Ausbreitungsmedium ruht. Bewegt sie sich dagegen, so sendet sie in Vorwärts- und Rückwärtsrichtung bezüglich ihrer Bewegungsrichtung Wellen mit unterschiedlicher Wellenlänge aus. Das wird besonders deutlich, wenn die Quelle in regelmäßigen Zeitabständen kurze Impulse emittiert, die sich dann mit unveränderter Geschwindigkeit im Medium ausbreiten. Die Quelle läuft diesen Impulsen entweder hinterher, was zur Verkürzung des räumlichen Impulsabstands (bzw. der Wellenlänge); oder von ihnen weg, was zur Vergrößerung des Impulsabstands führt.

2.3.2. Interferenz und Beugung

2.3.2.1. Interferenz. Mehrere Einzelwellen des gleichen Typs, beispielsweise Lichtwellen, können sich häufig ohne jede gegenseitige Beeinflussung überlagern, sich gewissermaßen gleichzeitig im gleichen Raumbereich bzw. Ausbreitungsmedium ausbreiten. Die *ungestörte Überlagerung* wird erst dann gestört, wenn nichtlineare Effekte auftreten. Bei elektromagnetischen Wellen im Vakuum konnten solche bis heute nicht gefunden werden, während sie beim Schall auftreten müssen. In einem durch eine Schallwelle veränderten Medium wird sich eine zweite Welle etwas anders verhalten als in der ungestörten Substanz.

Bei der Überlagerung sind die Feldgrößen der Einzelwellen unter Beachtung ihres Vorzeichens (bzw. ihres komplexen Phasenfaktors bei komplexwertigen Feldern und/oder ihrer Richtung bei Vektorfeldern) zu addieren. Dabei kompensieren sich positive und negative Beiträge von den Einzelwellen ganz oder teilweise. So löschen sich zwei harmonische ebene Wellen gleicher Amplitude, Frequenz und Wellenlänge gegenseitig völlig aus, wenn sie räumlich um eine halbe Wellenlänge bzw. zeitlich um eine halbe Schwingungsdauer gegeneinander verschoben sind. Ihre Phasen (↑ 2.3.1.3.) unterscheiden sich dann gerade um π. Als zweites Beispiel kann die Schwebungserscheinung (↑ Abb. 2.3.1-2) dienen: Bei gleicher Amplitude der beiden überlagerten Wellen gibt es zu jeder Zeit Punkte auf der x-Achse, d. h. Ebenen im Raum, auf denen das resultierende Feld verschwindet.

Die Überlagerung der Feldgrößen ist ein sehr einfaches Schema. Es führt dennoch zu etwas Besonderem, weil im allgemeinen die Feldgrößen selbst gar nicht gemessen werden, sondern die Intensitäten. Im wesentlichen ist die *Intensität* der zeitliche Mittelwert von u^2 (bzw. von $|u|^2 = uu^*$) der von Ort und Zeit abhängigen Feldgröße $u(r, t)$. Die zeitliche Mittelung über Zeiträume, die groß sind gegenüber der Schwingungsdauer der Welle, erfolgt meist „automatisch" durch das begrenzte zeitliche Auflösungsvermögen der benutzten Strahlungsempfänger. Eine Fotoplatte hat beispielsweise wegen der relativ großen Belichtungsdauer ein sehr geringes zeitliches Auflösungsvermögen.

Überlagern sie zwei Wellen u_1 und u_2 zu einem resultierenden Feld $u = u_1 + u_2$, so ist die entstehende Intensität $I = \overline{(u_1 + u_2)^2} = \overline{u_1^2} + \overline{u_2^2} + \overline{2u_1u_2} = I_1 + I_2 + \overline{2u_1u_2}$ im allgemeinen *nicht* gleich der Summe der Teilintensitäten. Der Querstrich deutet die zeitliche Mittelung an. Im Extremfall kann $I = 0$ sein an der Stelle, wo I_1 und I_2 maximale Werte annehmen. Die Gesamtintensität kann eine ausgeprägte räumliche Abhängigkeit besitzen (Interferenzstreifen u. ä.), obwohl I_1 und I_2 in dem betrachteten Raumbereich konstant sind. Es sind vor allem solche charakteristischen Effekte im Intensitätsverlauf, die als *Interferenzerscheinungen* bezeichnet werden. Das Auftreten von Interferenzen ist ein sicherer Beweis für den Wellencharakter der untersuchten Erscheinung. Es ist sehr wichtig, daß dieser Wellencharakter so durch alleinige Intensitätsmessungen erschlossen werden kann, weil die Feldgrößen selbst nur in seltenen Fällen direkt meßbar sind; bei Materiewellen sind sie sogar prinzipiell nicht meßbar.

Überlagern sich zwei ebene harmonische Wellen gleicher Wellenlänge mit den Amplituden A_1

und A_2 und einer Phasendifferenz α, also $u_1 = A_1 \cos(kx - \omega t)$ und entsprechend $u_2 = A_2 \cos(kx - \omega t + \alpha)$, so gilt $I_1 = A_1^2/2$ und $I_2 = A_2^2/2$, weil der Mittelwert der symmetrisch zwischen 0 und 1 hin und her pendelnden Funktion $\cos^2(\ldots)$ gerade 1/2 ist. Die Gesamtintensität ist

$$I = I_1 + I_2 + A_1 A_2 \left\{\overline{\cos\alpha + \cos[2(kx - \omega t) + \alpha]}\right\},$$

wobei das Produkt der beiden Kosinusfunktionen nach dem Additionstheorem umgeformt wurde. Der Mittelwert von $\cos[\ldots]$ ist Null; übrig bleibt $I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\alpha$.

Für $\alpha = 0$ sind beide Wellen „in Phase" und die Intensität erreicht ihren *Maximalwert* $(\sqrt{I_1} + \sqrt{I_2})^2$ entsprechend der Addition der Amplituden. Für $\alpha = \pi$ sind beide Wellen „gegenphasig" und schwächen sich maximal, die Intensität nimmt ihren *Minimalwert* $(\sqrt{I_1} - \sqrt{I_2})^2$ entsprechend der Subtraktion der Amplituden an. Zwischen beiden Extremfällen können alle Zwischenstufen auftreten. Die Phasendifferenz zwischen beiden Wellen beeinflußt wesentlich die resultierende Intensität.

Zu der bei Überlagerung zweier Wellen mit geringem Unterschied der Wellenlänge und Frequenz entstehenden *Schwebung* (↑ Abb. 2.3.1-2) gehört bei Mittelung über die raschen Oszillationen ein Intensitätsverlauf, der dem Quadrat der Amplitudenfunktion proportional ist und sich genau wie diese langsam verschiebt. Es entstehen also keine stationären, d. h. im Raum stillstehenden Interferenzverläufe. Falls das zeitliche Auflösungsvermögen der Empfänger ausreicht, kann man die *instationäre Interferenzerscheinung* noch beobachten. Falls nicht, d. h. falls auch noch über die wandernden Schwebungsmaxima gemittelt wird, ist die beobachtete Intensität konstant und gleich der Summe der Einzelintensitäten. Immer dann, wenn sich die Intensitäten addieren, also keine beobachtbaren Interferenzen entstehen, nennt man die Überlagerung der Wellen *inkohärent*; Voraussetzung für beobachtbare Interferenzen ist Kohärenz (↑ 2.3.2.2.) der Einzelwellen.

2.3.2.2. Kohärenz. Damit stationäre Interferenzfiguren entstehen, müssen die überlagerten Wellen gleiche Frequenz und Wellenlänge haben, anderenfalls sind die Wellen inkohärent (↑ 2.3.2.1.). Werden nicht nur zwei, sondern sehr viele Wellen überlagert, so genügt es nicht, wenn diese alle die gleiche Wellenlänge haben. Notwendig ist auch, daß ihre Phasen, d. h. also die Parameter α in $\cos(kx - \omega t + \alpha)$, nicht regellos von Welle zu Welle schwanken, sondern daß viele Einzelwellen gleiche Phasen haben und insgesamt nur wenige, die Gesamtheit aller denkbaren Werte $0 \leqq \alpha < 2\pi$ nicht überdeckende Phasenparameter α vorkommen. Anderenfalls entspricht das Ergebnis für die Gesamtintensität einer Mittelung über α in dem im vorigen Abschnitt abgeleiteten Intensitätsausdruck; wegen $\overline{\cos\alpha} = 0$ folgt dort $I = I_1 + I_2$; dem entspricht auch im allgemeinen Fall die für *inkohärente Überlagerung* kennzeichnende Addition der einzelnen Intensitäten. Regellos schwankende Phasenparameter haben die gleiche Wirkung wie Frequenzunterschiede: Sie „zerstören" die Interferenz.

Elektromagnetische Strahlung, die von einer makroskopischen Quelle emittiert wird, setzt sich aus den Beiträgen von ungeheuer vielen einzelnen, unabhängig voneinander strahlenden Atomen (Molekülen, Kernen ...) zusammen. Im allgemeinen schwankt auch bei gleicher Frequenz und Wellenlänge die Phase des emittierten Lichts statistisch regellos von Atom zu Atom. Zum Interferenzbild trägt daher nur die Interferenz jeder von einem Atom erzeugten Lichtwelle mit sich selbst bei. Dadurch werden die Interferenzen sehr „lichtschwach". Das von verschiedenen makroskopischen Lichtquellen bzw. von verschiedenen Stellen einer solchen Quelle emittierte Licht ist nicht kohärent, also nicht interferenzfähig. Eine wichtige Ausnahme ist das *Laserlicht* (↑ 4.1.4.), da bei einem strahlenden Laser eine fast perfekte Phasenstabilisierung der einzelnen atomaren Strahlungsquellen erfolgt. Daher ist Laserlicht hervorragend zur Durchführung von Interferenzversuchen geeignet.

Auch die von einem einzigen Atom ausgehenden Wellen sind nicht unbegrenzt interferenzfähig. Ein einzelner Emissionsvorgang dauert nur eine endliche Zeit τ. In dieser Zeit legt das Licht die Strecke $c\tau$ zurück, die daher die Länge eines emittierten Wellenpakets ist. Diese Länge $c\tau$ heißt *Kohärenzlänge L*. Erzeugt man in Interferenzanordnungen Gangunterschiede, die größer als die Kohärenzlänge sind, so verschwinden die Interferenzen.

Instruktiv ist auch eine andere Erklärung der endlichen Kohärenzlänge. Ein Vorgang der Dauer τ hat eine Frequenzschärfe (↑ 2.1.6.2.) $\Delta\omega \approx \tau^{-1}$, entsprechend hat ein Wellenpaket der endlichen Länge L eine Unschärfe $\Delta k \approx L^{-1}$ in der Wellenzahl (↑ 2.3.1.6.). Erzeugt man einen Gangunterschied der Größe l, so entsteht eine Unbestimmtheit $\Delta k \cdot l \approx l/L$ in der Phase der Welle. Ist $l \sim L$ oder gar $l \gg L$, so ist die Phasenunbestimmtheit wesentlich, und die einzelnen „Bestandteile" (im Sinne der Fourierdarstellung) des Wellenpakets verlieren ihre Kohärenz, sind nicht mehr interferenzfähig.

2.3.2.3. Beugung an Gittern. Wichtige Interferenzerscheinungen treten auf, wenn eine ebene Primärwelle dazu führt, daß von *periodisch angeordneten Zentren* Wellen ausgehen. Das geschieht bei der Streuung der Welle an Objekten (z. B. Atomen), die durch die Primär-

welle zur Emission von Streuwellen angeregt werden. Es kann aber auch dadurch erreicht werden, daß die Welle auf einen für sie undurchdringlichen Schirm mit entsprechenden Öffnungen bzw. durchlässigen Stellen fällt, von denen dann Kugel- oder Zylinderwellen ausgehen. In der Lichtoptik werden diese „Öffnungen" erzeugt, indem man in Glasplatten parallele Furchen ritzt. Die unverletzten Zwischenstreifen wirken als durchlässige Spalte.

Die Elementarwellen interferieren miteinander, wobei sie sich nur in ganz bestimmten Raumrichtungen verstärken. Diese Richtungen können von der Ausbreitungsrichtung der Primärwelle völlig verschieden sein. Da man alle Abweichungen von der gradlinigen Ausbreitung der Welle als *Beugung* bezeichnet, spricht man z. B. von Elektronen- oder Neutronenbeugung oder der Beugung von Licht an künstlich hergestellten Gittern.

2.3.2.3.1. *Strichgitter.* Die von verschiedenen Spalten ausgehenden Wellen (Abb. 2.3.2-1) verstärken sich in solchen Richtungen, für die der Gangunterschied der von benachbarten Spalten ausgehenden Wellen ein ganzzahliges Vielfaches $n\lambda$ der Wellenlänge ist. Bei senkrechtem Einfall der Primärwelle ist der Gangunterschied $a \sin \alpha$; wobei a die Gitterkonstante ist. Die genauere Theorie zeigt, daß die Schärfe der damit festgelegten Intensitätsmaxima mit der Anzahl M der Spalte anwächst. Darauf beruht die Möglichkeit, *hochauflösende Gitterspektrographen* zu bauen, bei denen die zu nur wenig verschiedenen Wellenlängen gehörenden Maxima noch getrennt werden. Das *Auflösungsvermögen* $\lambda/\Delta\lambda = Mn$ wird als Verhältnis von Wellenlänge zu der noch nachweisbaren Wellenlängendifferenz $\Delta\lambda$ definiert; n ist die ganze Zahl, die angibt, ob das erste, zweite usw. Intensitätsmaximum zur Messung benutzt wird. Man spricht dann von Interferenzen n-ter Ordnung. Mit Zunahme von n werden die Interferenzstreifen immer lichtschwächer. Für $M = 100000$ und $n = 3$ erhält man bereits ein Auflösungsvermögen von 300000.

Für eine unendlich große, von parallelen Spalten durchzogene Ebene werden die Interferenzmaxima (in einigem Abstand von der Ebene) beliebig scharf. Das Wellenfeld hat dann eine sehr einfache Struktur, es besteht nämlich aus einer Überlagerung ebener Teilwellen, Index n, deren Wellenlänge mit der Primärwellenlänge übereinstimmt. Also gilt auch $k_n = k_0$ für die Beträge der entsprechenden Wellenzahlvektoren $\boldsymbol{k}_n$ (↑ 2.3.1.3.). Das gesamte Wellenfeld muß sich um den gleichen Faktor wie die Primärwelle ändern, wenn man von einem beliebigen Punkt

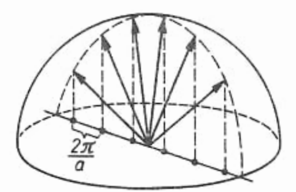


Abb. 2.3.2-2 Konstruktion der erlaubten Wellenvektoren bei der Beugung am Strichgitter; jeder der gezeichneten Wellenzahlvektoren kann zur Primärwelle gehören; die Kugel hat den Radius k_0

parallel zur Gitterebene weiterschreitet in Spaltrichtung bzw. um die Gitterkonstante a senkrecht zu den Spalten. Daraus folgt für die parallel zur Spaltebene orientierten Komponenten aller Wellenzahlvektoren im ersten Fall, daß die Komponente in Spaltrichtung mit dem entsprechenden Wert für die Primärwelle übereinstimmt; und im zweiten Fall, daß die Komponente senkrecht zur Spaltrichtung nur um ein ganzzahliges Vielfaches $(2\pi/a)\, n$ vom Primärwert abweichen kann. Ein solcher Unterschied bewirkt nämlich bei der Translation um a lediglich eine wirkungslose Phasenänderung um $2\pi n$. Für die Komponente des Wellenzahlvektors senkrecht zur Gitterebene folgen aus der Symmetrie des Problems keine Einschränkungen. Sie kann daher so gewählt werden, daß $k_n = k_0$ erfüllt ist (Abb. 2.3.2-2). Auf diese Weise können die Richtungen bestimmt werden, in denen sich hinter der Spaltebene ebene Wellen ausbreiten können.

2.3.2.3.2. *Netzebene.* Unter einer streuenden Netzebene soll eine zweifach periodische Anordnung von Streuzentren verstanden werden, von denen aufgrund der Anregung durch die Primärwelle Kugelwellen ausgehen. Die Behandlung des Strichgitters ist in zweierlei Hinsicht zu modifizieren: a) Es gibt auch „zurücklaufende Wellen", also Reflexe „vor" der Netzebene. b) Durch die Auflösung der Spalte in einzelne Punkte kann sich auch die Komponente von $\boldsymbol{k}$ in Spaltrichtung ändern. Bilden daher die Streuer ein rechtwinkliges Gitter, dessen Achsen zur Festlegung der x- und y-Richtung eines kartesischen Koordinatensystems benutzt werden, und sind a und b die

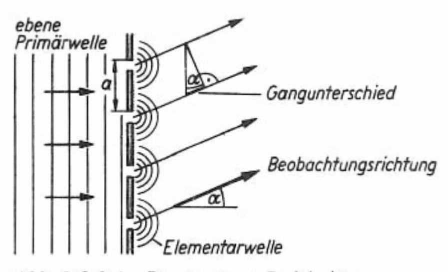


Abb. 2.3.2-1 Beugung am Strichgitter

entsprechenden Gitterkonstanten, so kann der Wellenzahlvektor in x-Richtung um Vielfache von $2\pi/a$ und in y-Richtung um Vielfache von $2\pi/b$ verändert werden. Die möglichen Vektoren $\boldsymbol{k}$ bilden also bei Projektion auf die x-y-Ebene ein Gitter mit den Gitterkonstanten $2\pi/a$ und $2\pi/b$. Beim Übergang zu einem nicht rechtwinkligen Gitter ändert sich qualitativ nichts; das Gitter der $\boldsymbol{k}$-Projektionen ist das um den Faktor 2π gedehnte zur Netzebene der Streuer reziproke Gitter. Die möglichen $\boldsymbol{k}$-Vektoren findet man durch eine zum Strichgitter (↑ Abb. 2.3.2-2) analoge Konstruktion unter Einbeziehung der „unteren" Halbkugel. Die angeregte Netzebene „strahlt" nach beiden Seiten unter wohldefinierten Richtungen.

Sobald die Wellenlänge λ groß gegen die Gitterkonstante ist, wird k_0 klein gegen $2\pi/a$ bzw. $2\pi/b$. Außer dem Primärstrahl $\boldsymbol{k}_0$ bleibt dann nur noch der genau nach dem Reflexionsgesetz reflektierte Strahl übrig. Die Netzebene verhält sich im *langwelligen Grenzfall* wie eine strukturlose Ebene. Umgekehrt kann man nur mit Wellenlängen, die von der Größenordnung der Gitterkonstanten sind, etwas über die Gitterstruktur erfahren. (Bei sehr kurzen Wellen wird die Situation durch die Vielzahl der gebeugten Strahlen sehr unübersichtlich.) Analoges gilt auch für die Ermittlung von Strukturen im räumlichen Fall.

2.3.2.3.3. *Raumgitter.* Beim Übergang von der Netzebene, also vom zweidimensionalen Gitter, zum dreidimensionalen Raumgitter bilden die Spitzen der erlaubten $\boldsymbol{k}$-Vektoren ein dreidimensionales Gitter. Es entfällt damit die Möglichkeit, eine Komponente des Wellenzahlvektors so zu wählen, daß sein Betrag den geforderten Wert k_0 erhält. Bei beliebiger Primärwelle, d. h. bei beliebigem $\boldsymbol{k}_0$, gibt es daher im allgemeinen überhaupt keine Raumrichtungen, unter denen ein starker abgebeugter Strahl entsteht. Insbesondere im langwelligen Grenzfall (↑ 2.3.2.3.2.), in dem sich das Gitter wie ein strukturloses Medium verhält, entstehen keinerlei „Reflexe". Nur dann, wenn gewissermaßen zufällig ein oder mehrere $\boldsymbol{k}$-Gitterpunkte auf der durch $k = k_0$ definierten sogenannten *Ewaldkugel* liegen, entsteht ein Reflex. Dementsprechend wird bei der Beugung an Raumgittern, die das entscheidende Hilfsmittel der Strukturanalyse darstellt, stets ein Parameter variiert, so die Wellenlänge oder die Orientierung des Kristalls.

Die Entstehung eines Reflexes bei der Beugung an einem Raumgitter kann so interpretiert werden, daß die an einer dem Reflex zugeordneten ausgezeichneten Schar zueinander paralleler Netzebenen reflektierten Wellen sich alle gleichphasig überlagern. Die so entstehenden Glanzwinkel genügen der Braggschen Gleichung (↑ 6.4.2.2.1.).

2.3.2.4. Streuung an einem begrenzten Objekt. Die bei der Wechselwirkung einer Welle mit einem streuenden Objekt entstehende Streuwelle kann man sich aufgebaut denken durch Überlagerung elementarer Kugelwellen, die von den einzelnen Punkten des Objekts ausgehen. Die Wellenlänge dieser Elementarwellen ist, so wie bei einer ebenen Welle, durch die (gegebene) Frequenz bestimmt. Die einzelnen Kugelwellen interferieren miteinander, wodurch auch in großem Abstand von dem streuenden Objekt eine komplizierte *Winkelabhängigkeit der Streuwelle* entstehen kann. Sind jedoch die linearen Abmessungen (in allen Raumrichtungen) des Streuens klein gegen die Wellenlänge, so gilt dies auch für die auftretenden maximalen Gangunterschiede. Die einzelnen Elementarwellen sind daher praktisch alle „in Phase", und es entsteht bei der Überlagerung eine völlig *isotrope Kugelwelle*, deren Amplitude gleich der Summe der Amplituden aller Elementarwellen ist. Das gesamte Objekt streut dann genau so wie ein *punktartiges Objekt*, von dem eine einzige Elementarwelle ausgeht. Aus der Streuung einer Welle kann also nicht auf Strukturen innerhalb des Streuers geschlossen werden, wenn die Wellenlänge wesentlich größer als dessen Abmessung ist.

2.3.3. Eigenschwingungen eines Felds

2.3.3.1. Schwingende Saite. Bei der Ausbreitung einer Welle in einem unbegrenzten Raumbereich sind prinzipiell alle Frequenzen und die zugehörigen Wellenlängen möglich. Dagegen sind in begrenzten Bereichen zeitlich periodische Abläufe nur mit bestimmten Frequenzen möglich, die man *Eigenfrequenzen* nennt. Ihre Entstehung soll am eindimensionalen Beispiel einer schwingenden Saite untersucht werden. Ist die Saite an einem Ende fest eingespannt, so muß an dieser Stelle die Schwingungsamplitude zu *allen* Zeiten verschwinden. Dies wird durch die mit einem Vorzeichenwechsel verbundene Reflexion eines beliebigen Signals (Abb. 2.3.3-1) am festen Ende erreicht. Bei einer harmonischen Welle führt die Überlagerung zu einer stehenden

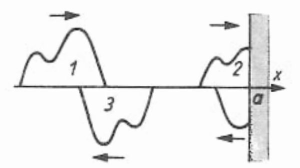


Abb. 2.3.3-1 Reflexion am eingespannten Ende; *1* bis *3* zeitlich aufeinanderfolgende Momentaufnahmen der Auslenkung

Welle, bei der die Auslenkung der schwingenden Saite proportional zu $\sin [k(x - a)] \cos (\omega t + \alpha)$ ist. $x = a$ bezeichnet die Stelle, an der die Saite eingespannt ist. Die Phasenkonstante α ist durch die Phasenlage der auf das feste Ende zulaufenden Welle bestimmt. Der Faktor $\sin [k(x - a)]$ sorgt dafür, daß das eingespannte Ende immer in Ruhe bleibt. Damit liegen aber überall die Stellen maximaler Schwingungsamplitude, also die durch $\sin [k(x_{max} - a)] = \pm 1$ bestimmten Schwingungsbäuche, und die durch die Nullstellen von $\sin [k(x - a)]$ bestimmten Schwingungsknoten fest (Abb. 2.3.3-2).

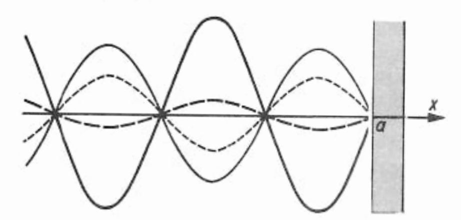


Abb. 2.3.3-2 Stehende periodische Welle; dargestellt ist die Auslenkung der schwingenden Saite für verschiedene Zeiten

Wird die Saite nun an zwei Stellen fest eingespannt, so muß sich vor beiden Enden eine stehende Welle ausbilden. Ausgehend von der Situation bei der einseitig eingespannten Saite (↑ Abb. 2.3.3-2) ist dies nur dann möglich, wenn das zweite feste Ende mit einem der ohnehin vorhandenen Schwingungsknoten zusammenfällt. Die Saitenlänge l ist dann stets ein ganzzahliges Vielfaches der halben Wellenlänge, also $\lambda/2$ für die *Grundschwingung*, λ für die erste *Oberschwingung* usw. Da die Wellenlänge $\lambda = 2\pi/k$ mit der Frequenz $\nu = v_{ph}/\lambda$ zusammenhängt, schwingt also die Saite bei gegebener Länge nur mit ganz bestimmten diskreten, für die Saite charakteristischen Eigenfrequenzen. Wegen $l = n\lambda/2$ mit $n = 1, 2, 3 \ldots$ gilt einfach $\nu_n = v_{ph}/\lambda_n = (v_{ph}/2l)\, n$. Die Grundfrequenz ist $v_{ph}/2l$, die Oberfrequenzen sind ganzzahlige Vielfache davon.

Diskrete Eigenfrequenzen treten auch auf, wenn ein oder beide Enden der Saite völlig frei schwingen können, d. h. wenn sich dort immer ein Schwingungsbauch befindet. Entscheidend ist der endliche Bereich, an dessen Grenzen die Feldfunktion bestimmte Randbedingungen erfüllen muß.

2.3.3.2. Verallgemeinerungen. Kompliziertere Systeme haben auch ein komplizierteres Spektrum von Eigenfrequenzen (Schwingungsspektrum) als die schwingende Saite. Bei zweidimensionalen (schwingende Platten und Membranen) oder dreidimensionalen Problemen kommt es auf die Form des begrenzten Bereichs an. Falls das Gebiet bestimmte Symmetrien hat, z. B. ein Würfel oder eine Kugel, können verschiedene Eigenschwingungen, die man oft als *Schwingungsmoden* oder kurz *Moden* bezeichnet, die gleiche Frequenz haben. Solche Moden mit gleicher Frequenz nennt man (miteinander) entartet. Dreidimensionale Eigenschwingungen des elektromagnetischen Felds werden in *Hohlraumresonatoren* erzeugt; das sind von (im betrachteten Frequenzbereich) sehr guten elektrischen Leitern begrenzte Hohlräume.

Es ist nicht erforderlich, daß der Bereich durch eine Oberfläche (bzw. Randkurve oder Randpunkte) scharf begrenzt ist. Es genügt für die Ausbildung diskreter Eigenfrequenzen, wenn das Wellenfeld außerhalb eines gewissen Bereichs verschwindet. Diese Situation ist für die Eigenschwingungen von Materiewellen, d. h. für die Lösungen der stationären Schrödingergleichung (↑ 2.5.4.1.) typisch.

Falls der Bereich nicht allseitig begrenzt ist, sondern gewissermaßen ein Loch hat, werden die Eigenschwingungen wegen der damit verbundenen Energieverluste gedämpft. So ist beispielsweise das schwingende Luftvolumen bei allen Blasinstrumenten nicht völlig abgeschlossen, sondern steht mit der „Außenluft" in Verbindung. An die Stelle einer scharf definierten Eigenfrequenz tritt bei solchen Problemen ein Resonanzgebiet, kurz eine *Resonanz*, deren Frequenzbreite (↑ 2.1.6.2.) durch die Größe der Dämpfung bestimmt wird. Dabei ist das Produkt $\tau\Delta\nu \approx 1$ aus Frequenzunschärfe und Lebensdauer τ einer angeregten und dann abklingenden Schwingung von der Größenordnung 1.

In der Elementarteilchenphysik machen sich hochgradig instabile Teilchen durch Resonanzen bemerkbar. In der normalen Quantenmechanik sind alle angeregten Zustände von Kernen, Atomen usw. eigentlich nur Resonanzzustände (↑ 2.5.6.3.). Will man diesen Umstand besonders betonen, so spricht man von zerfallenden oder quasistationären Zuständen. Zu all diesen Erscheinungen gibt es also analoge Effekte bei klassischen Wellenfeldern bzw. Eigenschwingungen.

2.4. Wahrscheinlichkeit und statistische Physik

2.4.1. Bedeutung und Bestimmung von Wahrscheinlichkeiten

2.4.1.1. Wahrscheinlichkeit und Häufigkeit. In den Sätzen der Umgangssprache „Morgen wird es wahrscheinlich regnen" oder „Es ist unwahrscheinlich, daß Herr Müller 100 Jahre alt wird", wird eine *Erwartung*, eine *Chance* mit einem der Wörter „wahrscheinlich" oder „un-

wahrscheinlich" charakterisiert. In einer exakten Wissenschaft müssen solche Aussagen über zu Erwartendes präzisiert, müssen Wahrscheinlichkeiten durch Zahlen quantitativ festgelegt werden. Für den Grenzfall, daß ein Ereignis unmöglich ist, soll die Wahrscheinlichkeit den Wert 0, und wenn es mit Sicherheit eintritt, den Wert 1 haben; allen dazwischenliegenden Fällen soll ein zwischen 0 und 1 liegender Zahlenwert der Wahrscheinlichkeit zugeordnet werden.

Wahrscheinlichkeitsaussagen über einmalige, nicht wiederholbare Vorgänge sind im Grunde sinnlos, weil sie durch das Experiment nicht als richtig oder falsch nachgewiesen werden können. Als Beispiel soll der Satz „Die Wahrscheinlichkeit, daß es morgen regnet, ist ungefähr 1/2" dienen. Tatsächlich wird es entweder regnen oder nicht regnen. In beiden Fällen wird die Aussage weder widerlegt noch bestätigt, da sie beide Möglichkeiten zuläßt. Eine solche Wahrscheinlichkeitsaussage über das Wetter kann nur begründet werden, wenn man den morgigen Tag als *Element einer Gesamtheit* von Tagen betrachtet, die hinsichtlich der Wetterchancen als gleichwertig angesehen werden, wobei sich für jeden dieser Tage entscheiden läßt, ob Regen fällt oder nicht. Der *Zahlenwert der Wahrscheinlichkeit* für Regen wird dann durch die *relative Häufigkeit* des Falls „Regen" in dieser „Versuchsreihe" bestimmt: Wahrscheinlichkeit für Regen = (Anzahl der Regentage)/(Gesamtzahl der Tage). Je nach der Auswahl der Gesamtheit von Tagen werden sich im allgemeinen verschiedene Wahrscheinlichkeiten ergeben. Beispiele: 1) Morgen ist der 15. April 1985 und ab 1886 hat es jeweils am 15. April 48mal geregnet und 52mal nicht. 2) Heute abend ist bei mäßigem Westwind der Himmel bedeckt und ab 1. 1. 1976 folgte auf einen solchen Tag 347mal Regen und 328mal kein Regen.

2.4.1.2. Statistische Kollektive. Das Wetterbeispiel zeigt eine subjektive Komponente des Wahrscheinlichkeitsbegriffs, die in der Willkür bei der Auswahl des Kollektivs liegt. Die Wetterprognose ist um so wertvoller, um so sicherer, je mehr Informationen und Gesetzmäßigkeiten bereits bei der Auswahl des Kollektivs dadurch berücksichtigt werden, daß man nur Tage zu einem Wetterkollektiv vereinigt, deren Vortag eine möglichst große Anzahl von Bedingungen erfüllt. Solche für das Wettergeschehen wesentlichen Parameter sind z. B. Temperatur, Luftdruck und -feuchtigkeit, Windstärke und -richtung. Um sie aufzufinden, benutzt man bereits erkannte Gesetzmäßigkeiten und schließt z. B. in der wissenschaftlichen Meteorologie die Geburtstage gewisser Heiliger als Wetterparameter aus.

Bei Wahrscheinlichkeitsaussagen über das Lebensalter kann das *Kollektiv eingeengt* werden *durch* Angaben über das bereits erreichte Alter, überstandene Krankheiten, Lebensgewohnheiten u. a. In jedem Fall verändern *neue Informationen* über Herrn Müller, dem die Prognose gilt, die Wahrscheinlichkeitswerte für die einzelnen Altersmöglichkeiten.

In vielen Fällen versucht man, die Anzahl der das Kollektiv kennzeichnenden Parameter soweit zu erhöhen, daß die Wahrscheinlichkeiten nur die Werte 0 oder 1 haben.

Bei den grundlegenden physikalischen Anwendungen des Wahrscheinlichkeitsbegriffs in der Statistischen Mechanik und in der Quantentheorie sind die statistischen Kollektive eindeutig fixiert.

2.4.1.3. Statistische Schwankungen. Die Festlegung von Wahrscheinlichkeiten durch Häufigkeiten ist nur dann für eine Prognose von Wert, wenn es sich um die Häufigkeit des betrachteten Ereignisses in einer hinreichend langen Versuchsreihe handelt. In kurzen Reihen z. B. aus nur 2, 3 oder 4 Versuchen, muß man damit rechnen, daß statistische Schwankungen einen falschen Wert der Wahrscheinlichkeit vortäuschen. Es ist z. B. ohne weiteres möglich und auch gar nicht besonders selten, daß sich beim dreimaligen Werfen einer Münze dreimal „Wappen" zeigt. Dennoch ist man davon überzeugt, daß „im Mittel" gleich oft Wappen wie Zahl fallen wird, und wird das auch in längeren Versuchsreihen bestätigt finden.

In einfachen Fällen, wie dem Werfen einer Münze oder eines Würfels, ist eine *unbegrenzte Fortsetzung der Versuchsreihe* zumindest denkbar. Man kann dann die Wahrscheinlichkeit als Grenzwert der relativen Häufigkeit in einer Reihe definieren, bei der die Gesamtzahl der Versuche nach Unendlich geht. Die bei endlichen Versuchsreihen zu erwartenden Abweichungen (Fluktuationen) zwischen der relativen Häufigkeit und ihrem Grenzwert Wahrscheinlichkeit können berechnet werden. Die über viele Versuchsreihen aus je N Versuchen gemittelten betragsmäßigen Abweichungen |Wahrscheinlichkeit minus relative Häufigkeit| sind proportional zu $1/\sqrt{N}$. Für $N \gg 1$ kann also eine weitgehende Übereinstimmung von Wahrscheinlichkeit und relativer Häufigkeit erreicht werden.

2.4.1.4. Schätzung und Berechnung von Wahrscheinlichkeiten. Bei einfachen Fragestellungen lassen sich Wahrscheinlichkeiten unter der Annahme voraus berechnen, daß verschiedene Fälle als gleichwahrscheinlich angesehen werden. Dafür kann es zwei Begründungen geben: a) *Symmetrieüberlegungen* und b) *mangelnde Kenntnis*. Beide Arten von Begründungen sind meist gleichzeitig anzutreffen. Der erste Fall liegt vor, wenn man die Wahrscheinlichkeit für „Wappen" beim Werfen einer Münze mit 1/2,

die für eine bestimmte Zahl beim Würfeln mit 1/6 ansetzt. Man betrachtet die beiden Seiten einer Münze oder die sechs Seiten des Würfels aus Symmetriegründen als gleichwertig, wenigstens so lange, bis eine genauere Untersuchung diese Annahme als nicht exakt erweist; z. B. wäre es denkbar, daß die mit den verschiedenen Prägungen auf beiden Seiten der Münze verbundenen Unsymmetrien in der Massenverteilung einen Einfluß auf das Versuchsergebnis haben. Die mangelnde Kenntnis wird durch eine Symmetrieannahme ersetzt. Im Falle des Würfels ist die Verschiebung zu dem zweiten Argument noch stärker. Der Würfel könnte gefälscht sein, ohne daß es ihm von außen anzusehen ist.

Durch *Reduktion* relativ komplizierter Situationen *auf gleichwahrscheinliche Fälle* können viele Aufgaben der Wahrscheinlichkeitstheorie gelöst werden. Mit den Mitteln der mathematischen *Kombinatorik* kann dann häufig berechnet werden, wieviele elementare gleichwahrscheinliche Fälle in einem komplexeren Fall der Fragestellung enthalten sind.

Wahrscheinlichkeiten, die auf diese Weise im voraus geschätzt bzw. berechnet werden, nennt man häufig a-priori-Wahrscheinlichkeiten; solche, die im Gegensatz dazu als Häufigkeiten in langen Versuchsreihen bestimmt werden, a-posteriori-Wahrscheinlichkeiten.

2.4.1.5. Zusammenfassung und Verallgemeinerung. Wahrscheinlichkeiten geben eine *quantitative Bewertung von Möglichkeiten.* Sie stellen Prognosen über mögliche Ereignisse dar, bei denen verschiedene Fälle eintreten können, und zwar bei jeder konkreten Realisierung nur ein einziger. Jedem Fall wird eine Wahrscheinlichkeit zugeordnet, die mit der Häufigkeit dieses Falls in einem möglichst großen Kollektiv von realisierten Ereignissen übereinstimmt. Dieses Kollektiv heißt *statistisches Kollektiv.* Es muß genau definiert werden, damit sinnvolle Wahrscheinlichkeitsaussagen gemacht werden können. Die Auswahl des Kollektivs hängt von den Informationen ab, die man über den jeweiligen Vorgang besitzt. Neue Informationen, die häufig in Vorentscheidungen oder Teilrealisierungen bestehen, führen zu einer *Reduktion des statistischen Kollektivs.*

2.4.1.6. Statistischer Entropiebegriff. Bei Vorgabe einer Wahrscheinlichkeitsverteilung, bei der verschiedene Fälle mit den ihnen zugeordneten Wahrscheinlichkeiten realisiert werden, ist eine exakte Vorhersage des bei einer konkreten Realisierung verwirklichten Falls nicht möglich. Vielmehr besteht eben eine gewisse Unsicherheit oder *Unbestimmtheit* bezüglich des eintretenden Ereignisses. Diese Ungewißheit ist jedoch bei verschiedenen Wahrscheinlichkeitsverteilungen nicht gleich groß. Beim Vergleich verschiedener Verteilungen (Abb. 2.4.1-1) wird im Fall a) der Fall mit der Nummer 2 mit Sicherheit verwirklicht, so daß die Unsicherheit verschwindet. Bei b) wird der Fall 2 noch mit überwiegender Wahrscheinlichkeit verwirklicht, aber durch die kleine Wahrscheinlichkeit w_3 entsteht eine geringe Unsicherheit über das zu erwartende Ergebnis. Diese ist bei c) bereits wesentlich größer, während bei d), also bei fünf gleichwahrscheinlichen Fällen, maximale Ungewißheit besteht.

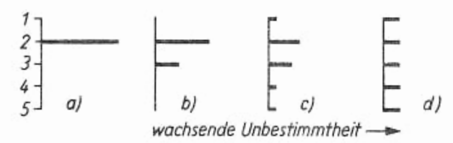


Abb. 2.4.1-1 Zum statistischen Entropiebegriff

Der statistische Entropiebegriff entsteht durch Verschärfung dieser qualitativen Diskussion. Dabei wird jeder diskreten Wahrscheinlichkeitsverteilung ein exakt definiertes, als *Entropie* bezeichnetes *Unbestimmtheitsmaß S* zugeordnet, das im obigen Beispiel (↑ Abb. 2.4.1-1) von a) bis d) monoton anwächst; bei a) ist S minimal und bei b) maximal. Dieses allgemeine Unbestimmtheitsmaß wurde zuerst von SHANNON im Zusammenhang mit statistischen Problemen der Informationsübertragung eingeführt. Der *allgemeine Entropieausdruck lautet*

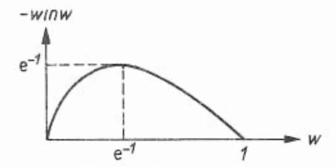


Abb. 2.4.1-2 Verlauf der Funktion $-w \ln w$

(1) $$S = -k \sum_n w_n \ln w_n,$$

wobei k zunächst eine beliebige positive Konstante sein kann. Diese Entropiedefinition sichert,

a) daß S niemals negativ ist und der Minimalwert $S = 0$ bei völliger Bestimmtheit (↑ Abb. 2.4.1-1 a) angenommen wird;

b) daß der Maximalwert $S_{max} = k \ln N$ bei Gleichverteilung (↑ Abb. 2.4.1-1 d) auf alle N Fälle auftritt;

c) daß die Wahrscheinlichkeiten aller Fälle nach dem gleichen Bildungsgesetz $-w_n \ln w_n$ (Abb. 2.4.1-2) zu S beitragen, die einzelnen Fälle also nicht etwa von vornherein als „wichtig" oder „unwichtig" bewertet werden. $w \ln w$ verschwindet sowohl für $w \to 0$ als auch für $w \to 1$;

d) daß sich die Entropie bei statistischer Unabhängigkeit zweier Merkmalgruppen (↑ 2.4.2.2.)

additiv aus den Entropien der beiden Einzelverteilungen zusammensetzt.

Je größer die Entropie einer Verteilung, um so geringer ist die durch Angabe dieser Verteilung vermittelte *Information*. Beispielsweise ist bei dem Wetterbeispiel (↑ 2.4.1.1.) die mit maximaler Entropie $S_{max} = k \ln 2$ verknüpfte Aussage: $w\{\text{Regen}\} = w\{\text{kein Regen}\} = 1/2$ völlig wertlos für die Entscheidung der Frage, ob man einen Regenschirm mitnehmen soll oder nicht. Die *thermodynamische Entropie* hat aufgrund des zweiten Hauptsatzes der Thermodynamik im thermodynamischen Gleichgewicht ihren maximalen Wert. Daher ist es einleuchtend, daß dieser Zustand durch ein Minimum von Informationen, die in der Angabe der wenigen thermodynamischen Parameter des Systems enthalten sind, vollständig gekennzeichnet werden kann.

2.4.2. Rechnen mit Wahrscheinlichkeiten

2.4.2.1. Addition von Wahrscheinlichkeiten und statistische Unabhängigkeit. Aus der Definition der Wahrscheinlichkeiten folgen einige einfache Eigenschaften und Rechenregeln. Es sollen N verschiedene Fälle durch einen von 1 bis N laufenden Index unterschieden werden. Im Wetterbeispiel oder beim Werfen einer Münze ist $N = 2$, beim Würfeln ist $N = 6$. Für ein bestimmtes Kollektiv sei der Satz der Wahrscheinlichkeiten gegeben. Gefragt wird nun nach der Wahrscheinlichkeit $w\{1 \text{ oder } 2\}$, daß *entweder* der Fall 1 mit w_1 *oder* der Fall 2 mit w_2 eintritt. Da sich die Häufigkeiten addieren, gilt $w\{1 \text{ oder } 2\} = w_1 + w_2$. An Stelle von 1 und 2 können beliebige andere Fälle treten. Die Wahrscheinlichkeit, daß überhaupt irgendein Fall eintritt, ist 1. Die w_n genügen mithin der *Normierungsbedingung* $\sum_n w_n = 1$. Die Wahrscheinlichkeit, daß der Fall n_0 nicht eintritt, daß also irgendeiner der $N - 1$ anderen Fälle verwirklicht wird, ist $1 - w_{n_0} = \sum_{n \neq n_0} w_n$. Hierbei erscheinen alle Fälle $n \neq n_0$ zu dem Fall „verschieden von n_0“ zusammengezogen.

2.4.2.2. Multiplikation von Wahrscheinlichkeiten. Die Fallunterscheidung nach dem Index n entspricht einer Sortierung der Ereignisse nach einem bestimmten Merkmal. Kommt den Ereignissen des gleichen Kollektivs noch ein weiteres Merkmal zu, das zu einer anderen Fallunterscheidung (Index $m = 1 \dots M$) führt, so existieren zunächst zwei Sätze w_n und w_m von Wahrscheinlichkeiten; für jeden einzelnen gelten die eben abgeleiteten Beziehungen. Nun kann man nach der Wahrscheinlichkeit w_{nm} fragen, daß *sowohl* der Fall n *als auch* der Fall m gleichzeitig eintreten. Die zugehörige Wahrscheinlichkeit kann im allgemeinen nicht aus den w_n und w_m bestimmt werden. Nur falls die Verteilung der n-Merkmale *statistisch unabhängig* von der Verteilung der m-Merkmale ist, gilt $w_{nm} = w_n w_m$. Darin kommt zum Ausdruck, daß die Häufigkeiten der m-Merkmale in dem Kollektiv, das durch Beschränkung auf die Ereignisse mit einem bestimmten Merkmal n entsteht, noch die gleichen sind, wie im Kollektiv aller Ereignisse. Beispiel: Ein Satz von Kugeln habe zwei Merkmalgruppen: leicht oder schwer ($n = 1, 2$) und rot, grün, gelb oder blau ($m = 1, 2, 3, 4$). Streicht man alle leichten Kugeln rot oder grün, alle schweren gelb oder blau an, so sind die Merkmale Gewicht und Farbe nicht unabhängig verteilt. Vielmehr besteht ein Zusammenhang, hier sogar eine feste Zuordnung. Es genügt aber bereits, die leichten Kugeln bevorzugt rot oder grün zu streichen, um die statistische Unabhängigkeit zu zerstören. In allen diesen Fällen spricht man von *Korrelation* der beiden Merkmale.

2.4.2.3. Mittelwert und Schwankung von Zufallsgrößen. Häufig können die zur Fallunterscheidung benutzten Merkmale selbst durch Zahlen charakterisiert werden. Dies ist typisch für naturwissenschaftlich-technische Anwendungen, bei denen jeder Fall durch bestimmte Werte physikalischer Größen gekennzeichnet wird; z. B. durch die Anzahl der Gasatome in einem bestimmten Raumgebiet oder durch die Geschwindigkeit eines Gasatoms. Die verschiedenen Werte einer solchen *Zufallsgröße* werden mit den zugeordneten Wahrscheinlichkeiten realisiert. Kann, wie im ersten Beispiel, die Merkmalsgröße nur diskrete Werte a_n annehmen, so können die Fälle wieder durch einen Index n unterschieden werden. Der *Mittelwert* der Größe ist $\bar{a} = \sum_n a_n w_n$; dabei ist jeder Wert mit seiner Wahrscheinlichkeit, also seiner Häufigkeit multipliziert. Sehr seltene Werte tragen fast gar nicht, häufige dagegen stark zu $\bar{a}$ bei. $\bar{a}$ selbst ist im allgemeinen gar kein möglicher Meßwert; z. B. ist eine Teilchenzahl stets ganzzahlig, die mittlere Teilchenzahl aber höchstens ausnahmsweise.

Neben dem Mittelwert $\bar{a}$ selbst braucht man ein Maß für die auftretenden *Abweichungen vom Mittelwert*, um eine Verteilung grob zu charakterisieren (Abb. 2.4.2-1). Dieses Maß ist die *mittlere quadratische Abweichung*, auch *Schwan-*

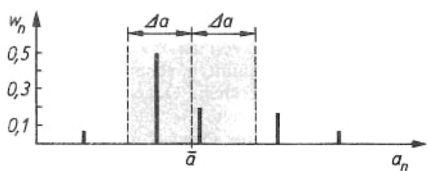


Abb. 2.4.2.-1 Zur Bedeutung von Mittelwert und Schwankung

kungsquadrat genannt, $\overline{(\Delta a)^2} = \overline{(a - \bar{a})^2} = \sum_n (a_n - \bar{a})^2 w_n$. Das Ausmultiplizieren des Quadrats in dieser Definitionsgleichung liefert

(2) $\overline{(\Delta a)^2} = \overline{a^2} - \bar{a}^2 = \sum_n a_n^2 w_n - \left(\sum_n a_n w_n\right)^2.$

Das Quadrieren kann dabei als einfache mathematische Möglichkeit angesehen werden, eine Kompensation positiver und negativer Abweichungen zu verhindern. Die Größe $\Delta a \equiv [\overline{(\Delta a)^2}]^{1/2}$ ist ein Maß für die Breite der Wahrscheinlichkeitsverteilung bzw. für die *Streuung der Meßwerte.*

Setzt sich die betrachtete Größe $A = \sum_i a_i$ additiv aus statistisch unabhängigen Teilgrößen a_i zusammen, so gilt dasselbe auch für das Schwankungsquadrat $\overline{(\Delta A)^2} = \sum_i \overline{(\Delta a_i)^2}$. Ist N die Zahl der Summanden, so gilt also $A \sim N$ und $\overline{(\Delta A)^2} \sim N$. Daraus folgt, daß die relativen Schwankungen $\Delta A/A = [\overline{(\Delta A)^2}]^{1/2}/A \sim 1/\sqrt{N}$ mit wachsender Zahl N der Summanden beliebig klein gemacht werden können. Dann können die Abweichungen vom Mittelwert im allgemeinen vernachlässigt werden und die Zufallsgröße verhält sich praktisch genau so wie eine Größe mit genau bestimmtem Wert. Dieser Sachverhalt ist entscheidend für die Beschreibbarkeit makroskopischer physikalischer Systeme mit den relativ einfachen Methoden der Thermodynamik, in der man von den Schwankungen, etwa des Energieinhalts eines Körpers, völlig absieht. Bei solchen Anwendungen erhält N die Bedeutung der Teilchenzahl des betrachteten Systems.

2.4.2.4. Zufallsgrößen mit kontinuierlichem Wertevorrat. Die Geschwindigkeit eines Gasatoms kann jeden beliebigen positiven Wert annehmen; sie durchläuft einen kontinuierlichen Wertebereich. In diesem und in allen ähnlichen Fällen ist es sinnlos, von der Wahrscheinlichkeit für einen ganz bestimmten, scharfen Geschwindigkeitswert zu sprechen. Dagegen kann man die Wahrscheinlichkeit Δw dafür betrachten, daß ein Gasatom eine Geschwindigkeit in dem zwischen v und $v + \Delta v$ gelegenen Intervall hat. Für kleine Δv ist Δw zu Δv proportional. Der Grenzwert von $\Delta w/\Delta v$ für verschwindende Intervallbreite, d. h. für $\Delta v \to 0$, ist die *Wahrscheinlichkeitsdichte* $w(v)$. Die Wahrscheinlichkeit, eine Geschwindigkeit aus dem endlichen Intervall $v_1 \leqq v \leqq v_2$ zu messen, erhält man durch Summieren der Teilintervalle bzw. Integration: $w\{v_1 \dots v_2\} = \int_{v_1}^{v_2} \mathrm{d}v\, w(v)$.

Dementsprechend sind in den Gleichungen für diskrete Fälle, insbesondere in denen für Normierung, Mittelwert und Schwankungsquadrat die Summationen über einzelne Wahrscheinlichkeiten w_n durch Integration über die Wahrscheinlichkeitsdichte zu ersetzen. Integriert wird dabei über den gesamten Wertebereich, im Fall des Geschwindigkeitsbetrags also von Null bis Unendlich. Ort und Impuls eines Teilchens sind weitere wichtige Beispiele für Größen mit kontinuierlichem Wertevorrat.

2.4.3. Einfache physikalische Anwendungen

2.4.3.1. Teilchenzahlschwankungen und Binomialverteilung. Ein ideales Gas bestehe aus N Teilchen im Volumen V. Betrachtet man ein beliebiges Teilvolumen V', so ist $p = V'/V$ die Wahrscheinlichkeit, ein bestimmtes, aber beliebig herausgegriffenes Teilchen in V' zu finden. Wird die Größe V' des Teilvolumens verändert, so verkleinert oder vergrößert sich p im gleichen Maße wie V. Offenbar ist $1 - p$ die Wahrscheinlichkeit, das betrachtete Teilchen *nicht* in V' zu finden. Die Wahrscheinlichkeit, *n bestimmte* Teilchen in V' und die restlichen $N - n$ außerhalb V' anzutreffen, ist $p^n(1-p)^{N-n}$, das ist also das Produkt aus den insgesamt N Faktoren p bzw. $(1 - p)$ für die jeweils einem einzelnen Teilchen zugeordneten Wahrscheinlichkeiten. Dabei ist berücksichtigt, daß p für alle Teilchen den gleichen Wert hat und daß die Teilchen wegen der im idealen Gas fehlenden Teilchenwechselwirkung statistisch unabhängig voneinander sind. Die Wahrscheinlichkeit, ein bestimmtes Teilchen in V' zu finden, hängt also nicht davon ab, ob sich schon andere Teilchen dort aufhalten.

Im allgemeinen interessiert man sich nur für Fragen, bei denen die einzelnen Teilchen nicht voneinander unterschieden werden; die Teilchen

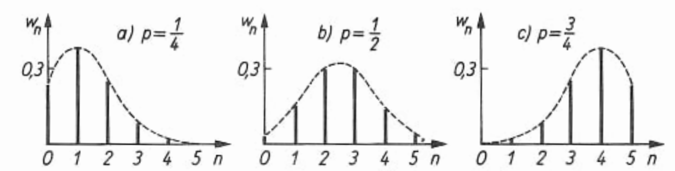


Abb. 2.4.3-1 Binomialverteilung für $N = 10$ und drei verschiedene p-Werte

bleiben „anonym“, können gegeneinander ausgetauscht werden. Eine typische Fragestellung ist: Mit welcher Wahrscheinlichkeit w_n befinden sich genau $n \leqq N$ Teilchen in V', unabhängig davon, um welche es sich dabei handelt. Da es aufgrund relativ elementarer kombinatorischer Überlegungen

$$\binom{N}{n} = \frac{N(N-1)\dots(N-m+1)}{1\cdot 2\dots n} = \frac{N!}{(N-n)!\,n!}$$

verschiedene Möglichkeiten gibt, aus den insgesamt N Teilchen n auszuwählen, folgt

$$(3) \quad w_n = \binom{N}{n} p^n (1-p)^{N-n}.$$

Die Koeffizienten $\binom{N}{n}$ heißen *Binomialkoeffizienten*, die Verteilung $\{w_n\}$ dementsprechend *Binomialverteilung*. Für nicht zu große Werte von N ist die unmittelbare Auswertung von w_n problemlos (Abb. 2.4.3-1). Bei $p = 1/2$ ist $p = 1 - p$ und $w_n = \binom{N}{n} \Big/ 2^N$ unterscheidet sich von den Binomialkoeffizienten nur durch den (bei festem N) konstanten Faktor 2^{-N}. Dann gilt $w_n = w_{N-n}$, d. h., die Wahrscheinlichkeitsverteilung ändert sich nicht, wenn man die Zahl n der Teilchen in V' durch die Zahl $N - n$ der Teilchen außerhalb V' ersetzt.
Für die mittlere Teilchenzahl ergibt sich $\bar{n} = Np$. Dieses Ergebnis drückt einen sehr einfachen Sachverhalt aus: Im Mittel befinden sich $Np = NV'/V$ Teilchen in V', die mittlere Teilchenzahl ist proportional zum betrachteten Volumen, mit N/V als der mittleren Teilchendichte des Gases ist $\bar{n} = (N/V)\, V'$.
Das Schwankungsquadrat hat den Wert $\overline{(\Delta n)^2} = Np(1-p)$. Es wird klein für $p \ll 1$, weil sich dann überhaupt nur wenig Teilchen in dem kleinen Teilvolumen aufhalten, aber auch für $1 - p \ll 1$, weil sich dann nur sehr wenig außerhalb, also fast alle in V' befinden. Die relativen Teilchenzahlschwankungen sind von der Größenordnung $\Delta n/\bar{n} = [\overline{(\Delta n)^2}]^{1/2}/\bar{n} = \sqrt{(1-p)/\bar{n}}$, also umgekehrt proportional zu $\sqrt{\bar{n}}$. Meist ist $p \ll 1$. In 1 cm^3 eines Gases befinden sich bei Zimmertemperatur und Atmosphärendruck etwa 10^{20} Gasmoleküle, also ist $\Delta n/\bar{n} \approx 10^{-10}$. So kleine Schwankungen sind kaum nachweisbar.

2.4.3.2. Gauß- und Poissonverteilung. Meist ist $N \gg 1$. Ist auch $\bar{n} \gg 1$ und $N - \bar{n} \gg 1$, so ist die Verteilung $\{w_n\}$ so um ihr bei $\bar{n}$ liegendes Maximum konzentriert (Abb. 2.4.3-2), daß n-Werte in der Nähe der Grenzen 0 und N des möglichen n-Bereichs mit vernachlässigbar kleinen Wahrscheinlichkeiten auftreten. Bei der Auswertung kann daher $N \gg 1$, $n \gg 1$, $N - n \gg 1$ angenommen werden. Dies erlaubt die Anwendung der Stirlingschen Formel zur Auswertung der die Binomialkoeffizienten aufbauenden Fakultäten und führt auf die sogenannte *Gaußverteilung*

$$(4) \quad w_n = \frac{1}{2\pi\overline{(\Delta n)^2}} \exp\left[-\frac{(n-\bar{n})^2}{2\overline{(\Delta n)^2}}\right]$$

$$\text{mit} \quad \overline{(\Delta n)^2} = \bar{n}(1-p).$$

Ist zusätzlich zu den getroffenen Voraussetzungen auch noch $p \ll 1$, also $\bar{n} \ll N$, so kann $\overline{(\Delta n)^2} \approx \bar{n}$ durch die mittlere Teilchenzahl ersetzt werden.

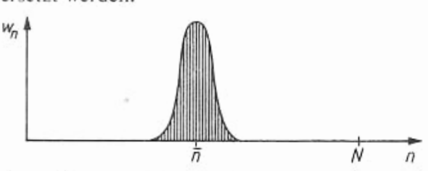


Abb. 2.4.3-2 Binomialverteilung für $1 \ll \bar{n} \ll N$

Im Falle $N \gg 1$, $p \ll 1$, d. h. falls n sehr klein gegenüber N bleibt, gilt die sogenannte *Poissonverteilung*

$$(5) \quad w_n = (\bar{n})^n \, e^{-\bar{n}}/n! \quad \text{mit} \quad \overline{(\Delta n)^2} = \bar{n},$$

wobei n selbst nicht groß zu sein braucht. Falls auch $\bar{n} \gg 1$ gilt, geht die Poissonverteilung in die Gaußverteilung über.
Die am Beispiel der Teilchenzahlschwankungen im idealen Gas diskutierte Binomialverteilung ist offenbar immer dann anwendbar, wenn es bei mehreren bzw. vielen unabhängigen Teilchen für jedes Teilchen zwei sich gegenseitig ausschließende Möglichkeiten mit den Wahrscheinlichkeiten p und $1 - p$ gibt. Im folgenden (↑ 2.4.3.3. bis 2.4.3.5.) werden dafür drei weitere physikalische Beispiele behandelt.

2.4.3.3. Irrflug. Für ein Teilchen soll es eine unbegrenzte diskrete Menge von möglichen räumlichen Lagen geben, z. B. Zwischengitterplätze bei der Diffusion von Atomen in einem kristallinen Festkörper. Bei einem elementaren „Sprung“ oder „Schritt“ geht das Teilchen in eine benachbarte Lage über. Nach insgesamt N Schritten kann es verschiedene Plätze s mit gewissen Wahrscheinlichkeiten w_s erreichen. Als einfaches Modell werde eine Kette möglicher Plätze betrachtet, wobei das Teilchen bei jedem Schritt mit der Wahrscheinlichkeit p nach rechts, d. h. von s nach $s + 1$, und mit der Wahrscheinlichkeit $1 - p$ nach links, d. h. von s nach $s - 1$, gelangt. Offenbar gibt die Binomialverteilung [↑ 2.4.(3)] die Wahrscheinlichkeit dafür an, daß von insgesamt N Einzelschritten n nach rechts erfolgen,

das Teilchen also von dem Anfangsplatz s_0 nach $s_0 + n - (N - n)$ gelangt. Startet das Teilchen bei $s_0 = 0$, so erreicht es den Platz $s = 2n - N$. Nach dem Ergebnis (↑ 2.4.3.1.) $\bar{n} = Np$ ist $\bar{s} = N(2p - 1)$, das ist positiv (negativ) für $p > 1/2$ ($p < 1/2$), d. h., falls das Teilchen bevorzugt nach rechts (links) springt. Die mittlere Änderung $\overline{\delta s}$ von s beim Einzelschritt ist $p(+1) + (1 - p)(-1) = 2p - 1$, bei N statistisch unabhängigen gleichartigen Schritten (↑ 2.4.2.3.) ist dies in Übereinstimmung mit dem obigen Resultat für $\bar{s}$ einfach mit der Schrittzahl N zu multiplizieren. Für $p = 1/2$ ist $\bar{s} = 0$, die Verteilung ist dann also symmetrisch um den Startpunkt 0, gleich große Wegstrecken nach beiden Seiten sind gleichwahrscheinlich.

Die Schwankungsbreite folgt zunächst wieder aus dem Zusammenhang zwischen s und n und dem bekannten Resultat für die Binomialverteilung (↑ 2.4.3.1.). Danach gilt $\overline{(\Delta s)^2} = \overline{\Delta(2n)^2} = 4\overline{(\Delta n)^2} = 4p(1 - p)\,N$.

Andererseits ist wegen der statistischen Unabhängigkeit der Einzelschritte (↑ 2.4.2.3.) $\overline{(\Delta s)^2}$ gerade das N-fache der entsprechenden Größe $\overline{(\delta s)^2} = p[+1 - \overline{\delta s}]^2 + (1 - p)[-1 - \overline{\delta s}]^2 = p(-2p + 2)^2 + (1 - p)(2p)^2 = 4p(1 - p)$ für den Einzelschritt. Die zugehörige mittlere Flugstrecke $\overline{\delta s} = 2p - 1$ wurde eben berechnet. $\Delta s \sim N$ ist ein Maß für die mittlere Entfernung eines Teilchens vom Ausgangspunkt nach N Sprüngen. Für den durch $p = 1/2$ charakterisierten symmetrischen Irrflugprozeß ist $\Delta s = \sqrt{N}$, das ist der bei fester Schrittzahl N mögliche Maximalwert. Jede Abweichung von der Symmetrie führt zu $\bar{s} \neq 0$, d. h. zu einer Fortbewegung auch im Mittel, und gleichzeitig zu einer Verminderung der Schwankungsbreite. Schließlich muß diese ja für $p = 0$ oder $p = 1$, also bei einer fixierten Sprungrichtung, verschwinden.

2.4.3.4. Exponentielles Zerfallsgesetz. Beim radioaktiven Zerfall von Atomkernen (↑ 1.1.5.3., 3.4.3.1.), bei der spontanen Emission von Lichtquanten durch atomare Systeme (↑ 1.2.2.2.) und bei anderen Umwandlungsprozessen erfolgen die Zerfallsprozesse der Einzelobjekte völlig unabhängig voneinander. Als konkretes Beispiel soll die Emission von α-Teilchen durch radioaktive Kerne gewählt werden. Die emittierten α-Teilchen werden mit geeigneten Zähleinrichtungen registriert. Bei einem hinreichend schwachen Präparat können die Einzelprozesse zeitlich getrennt werden.

Das Verhalten eines Einzelkerns hängt überhaupt nicht davon ab, ob die anderen Kerne des Präparats bereits zerfallen sind oder nicht, der betrachtete Kern wird also durch den Zerfall vieler „Bruderkerne" weder stabilisiert noch zum Zerfall angeregt. Der Zeitpunkt des Zerfalls ist völlig zufällig. Ein noch nicht zerfallener Kern wird im jeweils nächsten kleinen Zeitintervall Δt mit einer zur Intervalldauer proportionalen Wahrscheinlichkeit $\gamma\,\Delta t$ zerfallen. Dabei hängt γ nicht von der Zeit, also von der bisherigen Lebensdauer des Kerns ab. Im Gegensatz etwa zu höheren Lebewesen altern die Kerne nicht, es treibt sie kein verborgener Mechanismus zum Zerfall.

Um die *Überlebenswahrscheinlichkeit* $p(t)$ für eine endliche Zeitdauer 0 bis t zu finden, wird diese Zeit in eine große Anzahl M kleiner Zeitintervalle Δt eingeteilt. Die Überlebenswahrscheinlichkeit für ein Intervall ist $1 - \gamma\,\Delta t$. Nun soll der Kern sowohl im ersten wie im zweiten ... wie im M-ten Intervall überleben. Die zugehörige Wahrscheinlichkeit ist gleich dem Produkt $(1 - \gamma\,\Delta t)^M \to \exp(-\gamma t)$ (↑ 2.4.2.2.) aller Einzelwahrscheinlichkeiten für die M Intervalle. Der Pfeil zeigt den Grenzübergang $M \to \infty$, $\Delta t \to 0$ an, wobei $t = M\,\Delta t$ konstant bleibt. Nach diesem *exponentiellen Zerfallsgesetz* $p(t) = \exp(-\gamma t)$ gilt $p(t) \simeq 0$ für $\gamma t \gg 1$. In letzterem Falle zerfällt ein beliebiger aus der Menge aller Kerne herausgegriffener Kern innerhalb der Zeit t fast mit Sicherheit. Die Wahrscheinlichkeit dafür, daß ein Kern zur Zeit t überlebt und in dem anschließenden Intervall dt zerfällt, ist $p(t)\,\gamma\,dt$. Die *mittlere Lebensdauer* ergibt sich durch Mittelung über alle Intervalle als $\tau = \gamma \int_0^\infty dt\, e^{-\gamma t} t = \frac{1}{\gamma}$. Daher kann das Zerfallsgesetz auch in der Form $p(t) = \exp(-t/\tau)$ dargestellt werden. Nach der Halbwertszeit $\tau \ln 2$ ist $p = 1/2$, d. h., es ist (im Mittel) die Hälfte der Kerne zerfallen.

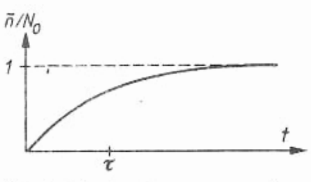


Abb. 2.4.3-3 Zum exponentiellen Zerfallsgesetz

Enthält ein radioaktives Präparat anfangs N_0 instabile Kerne, so ist

$$(6) \qquad \bar{N}(t) = N_0 p(t) = N_0\, e^{-\gamma t}$$

die mittlere Zahl der zur Zeit t noch nicht zerfallenen Kerne. Entsprechend ist $\bar{n}(t) = N_0[1 - p(t)]$ die mittlere Zahl der Zerfälle (Abb. 2.4.3-3), also der registrierten α-Teilchen. Solange $\gamma t \ll 1$, also $t \ll \tau$ bleibt, ist $\bar{n} \ll N_0$, es zerfällt also nur ein kleiner Teil der insgesamt vorhandenen Kerne. Registriert man die Zahl

$n(t)$ viele Male, was wegen der praktischen Konstanz der Gesamtzahl aller Kerne am gleichen Präparat in aufeinanderfolgenden Zeiten gleicher Dauer t erfolgen kann, so schwankt $n(t)$ um $\bar{n}(t)$ entsprechend der für diese Situation gültigen Poissonverteilung [↑ 2.4.(5)].

2.4.3.5. Wirkungsquerschnitt und mittlere freie Weglänge. Trifft ein Teilchenstrahl auf ein Hindernis, das klein gegenüber dem Strahlquerschnitt ist, so ist die Zahl der mit dem Hindernis wechselwirkenden Teilchen proportional zur *Stromdichte* (s. a. 2.2.2.2.) $j_0 = (\mathrm{d}N/\mathrm{d}A\mathrm{d}t)_0$ im ungestörten Strahl. Dabei ist $\mathrm{d}N$ die Zahl von Teilchen, die innerhalb einer kleinen Zeit $\mathrm{d}t$ eine kleine Fläche $\mathrm{d}A$ passieren. Die Wechselwirkung kann beispielsweise zur elastischen oder

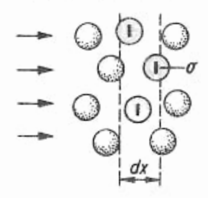


Abb. 2.4.3-4 Zur Wechselwirkung mit einer Schicht von Hindernissen; die Hindernisfläche ist nur für die grauen, der betrachteten Schichtdicke $\mathrm{d}x$ zugeordneten Atome eingezeichnet

unelastischen Streuung oder auch zur Absorption des betreffenden Teilchens führen. Der Quotient $\sigma = \dfrac{(\mathrm{d}N/\mathrm{d}t)_{\mathrm{WW}}}{(\mathrm{d}N/\mathrm{d}A\,\mathrm{d}t)_0} = \dfrac{(\mathrm{d}N/\mathrm{d}t)_{\mathrm{WW}}}{j_0}$ ist unabhängig von der Strahlintensität und charakterisiert daher das Hindernis. Im Zähler steht die Zahl der Teilchen, die mit dem Hindernis wechselwirken, also von ihm beeinflußt werden. σ hat die Dimension einer Fläche und wird *Wirkungsquerschnitt* genannt, bei der Betrachtung bestimmter Prozesse Wirkungsquerschnitt für Absproption bzw. Streuung oder kurz Absorptions- bzw. Streuquerschnitt. In der Darstellung $(\mathrm{d}N/\mathrm{d}t)_{\mathrm{WW}} = \sigma j_0$ wird deutlich, daß man σ als eine *effektive Hindernisfläche* interpretieren kann, in dem Sinne, daß jedes auf diese Fläche σ auftreffende Teilchen beeinflußt wird, während alle anderen ungestört weiterfliegen. Bei der Absorption durch einen schwarzen Körper würde σ_{abs} mit dessen geometrischem Querschnitt σ_{geom} übereinstimmen, während bei einem teilweise durchlässigen Hindernis $\sigma_{\mathrm{abs}} < \sigma_{\mathrm{geom}}$ wäre. σ hängt im allgemeinen von der Teilchenenergie ab.

Häufig wird ein einzelnes „Hindernis" durch ein Atom, einen Atomkern oder andere in großer Zahl vorhandene Objekte gebildet. Sind diese Objekte mit der Dichte n = Zahl der „Atome"/Volumen im Mittel gleichmäßig im Raum verteilt, so wächst die Wahrscheinlichkeit dafür, daß ein hineingeschossenes Teilchen mit einem dieser „Atome" in Wechselwirkung tritt, mit der Dicke der *durchstrahlten Schicht.* Betrachtet man zunächst eine dünne Schicht der Dicke $\mathrm{d}x$ und der (beliebig großen, aber insgesamt bestrahlten) Fläche A, so erhält diese die Mittelpunkte von $(\mathrm{d}Z)^{\mathrm{Schicht}} = nA\,\mathrm{d}x$ Atomen (Abb. 2.4.3-4). Jedem dieser Punkte ist eine Hindernisfläche σ zugeordnet. Die Wahrscheinlichkeit, daß ein beliebiges aus einem die gesamte Schicht treffenden Teilchenstrahl herausgegriffenes Teilchen mit irgendeinem dieser Atome „reagiert", ist $\mathrm{d}w = \dfrac{(\mathrm{d}N/\mathrm{d}t)^{\mathrm{Schicht}}_{\mathrm{WW}}}{j_0 A}$, also gleich dem Quotienten aus der Zahl der innerhalb $\mathrm{d}t$ in der Schicht beeinflußten Teilchen und der Gesamtzahl der im Strahl innerhalb $\mathrm{d}t$ auf die Schicht auftreffenden Teilchen. Da sich in einer hinreichend dünnen Schicht die einzelnen Hindernisse nicht gegenseitig verdecken, ist $(\mathrm{d}N/\mathrm{d}t)^{\mathrm{Schicht}}_{\mathrm{WW}} = (\mathrm{d}N/\mathrm{d}t)_{\mathrm{WW}}\,(\mathrm{d}Z)^{\mathrm{Schicht}}$. Mit $(\mathrm{d}N/\mathrm{d}t)_{\mathrm{WW}} = j_0\sigma$ und dem oben angegebenen Wert für $(\mathrm{d}Z)^{\mathrm{Schicht}}$ entsteht $\mathrm{d}w = n\sigma\,\mathrm{d}x$. Die Wahrscheinlichkeit, daß mit einem die dünne Schicht passierenden Teilchen etwas geschieht, ist danach proportional zur Schichtdicke $\mathrm{d}x$, zur Dichte n, mit der die störenden Zentren verteilt sind, und zum Wirkungsquerschnitt σ eines einzelnen Atoms.

Die mathematische Aussage $\mathrm{d}w = n\sigma\,\mathrm{d}x$ ist völlig analog zu $\mathrm{d}w = \gamma\mathrm{d}t$ beim radioaktiven Zerfall (↑ 2.4.3.3.). Also können die dortigen Überlegungen sinngemäß übertragen werden: Analog zur Überlebenswahrscheinlichkeit $p(t)$ ist $w(x) = \exp(-n\sigma x)$ die *Durchdringungswahrscheinlichkeit* bei einer endlichen Schichtdicke x. Analog zur mittleren Lebensdauer $\bar{t} = \gamma^{-1}$ ist $\bar{x} = (n\sigma)^{-1}$ die mittlere Strecke, die ein Teilchen im Medium bis zur Reaktion mit einem der Atome ungestört zurücklegt. Diese wichtige Größe heißt *mittlere freie Weglänge* und wird meist mit l bezeichnet, also $l^{-1} = n\sigma$.

2.4.4. Gleichgewichtsstatistik

2.4.4.1. Mikrozustände und ihre Wahrscheinlichkeiten. Jedes physikalische System kann sich in mehreren, oft sogar in sehr bzw. unendlich vielen Zuständen befinden. Das gilt unabhängig davon, ob das jeweilige System klassisch oder quantenmechanisch beschrieben wird, ob es sich also um den Zustandsbegriff der klassischen oder den der Quantenmechanik handelt. Falls das System selbst makroskopisch ist, so spricht man oft von seinen Mikrozuständen, um diese so dem phänomenologischen, auf die innere Struktur des Systems überhaupt nicht Bezug nehmenden Zustandsbegriff der thermodyna-

mischen Beschreibungsweise (↑ 6.1.1.2.) gegenüberzustellen. Es ist eine Aufgabe der klassischen bzw. der Quantenmechanik, die Mikrozustände selbst und die Vorgänge beim Übergang von einem zum anderen näher zu untersuchen. Falls aber die Bedingungen für die zeitliche Entwicklung des Systems nicht völlig genau bestimmt sind, sondern als Folge der schwachen *Wechselwirkung des Systems mit seiner Umgebung* ein statistisches Element enthalten, so müssen mechanische und statistische Methoden kombiniert werden, um die Wahrscheinlichkeiten zu bestimmen, mit denen ein System seine verschiedenen Mikrozustände annimmt. Die Untersuchung solcher Probleme bildet den Gegenstand der Statistischen Mechanik. Die Antwort auf die gestellte Frage ist überraschend einfach und allgemein (↑ 2.4.4.2.), falls sich das System in sich und mit seiner Umgebung im thermodynamischen Gleichgewicht befindet, d. h., nachdem die Besonderheiten der Anfangssituation durch irreversible Ausgleichsvorgänge vollständig verwischt, aus dem Systemverhalten eliminiert worden sind.

Die Wahrscheinlichkeiten sind relative Häufigkeiten in einem statistischen Kollektiv (↑ 2.4.1.2.) das hier durch eine Gesamtheit, ein sogenanntes Ensemble gleichartiger physikalischer Systeme gebildet wird. Dieses Kollektiv kann man sich auf zweierlei Art entstanden denken: durch gleichzeitige Untersuchung sehr vieler gleich gebauter und gleich vorbehandelter Systeme oder durch Untersuchung eines Systems zu sehr vielen (hinreichend weit auseinanderliegenden) verschiedenen Zeiten.

Die im thermodynamischen Gleichgewicht realisierten quantenmechanischen Zustände sind dessen *stationäre Zustände* (↑ 2.5.4.1.), in denen alle quantenmechanischen Wahrscheinlichkeiten und Mittelwerte zeitunabhängig sind. Diese stationären Zustände der vollentwickelten Quantentheorie entsprechen den „erlaubten" Quantenzuständen der älteren Bohrschen Theorie (↑ 1.1.4.5.), das System befindet sich also nach dieser vereinfachten Auffassung mit gewissen Wahrscheinlichkeiten in seinen verschiedenen *diskreten Quantenzuständen*, zwischen denen plötzliche Übergänge in Form sogenannter *Quantensprünge* erfolgen. Diese Zustände sind numerierbar und daher durch einen diskreten Index k charakterisierbar. Den Grundprinzipien der Quantenmechanik, insbesondere der Heisenbergschen Unbestimmtheitsrelation (↑ 2.5.1.7.) entsprechend, wird der Mikrozustand des Systems auf diese Weise vollständig und erschöpfend beschrieben, eine weitere Verfeinerung der Systemanalyse ist nicht möglich.

Die Quantenzustände eines Systems aus wechselwirkungsfreien und unterscheidbaren Teilchen werden beschrieben durch die Angabe der *Einteilchenzustände*, die jedes der numerierten Teilchen einnimmt. Die typischen und wichtigen Mehrteilchensysteme bestehen allerdings aus ununterscheidbaren, *identischen Teilchen* (↑ 2.5.5.3.), die prinzipiell nicht numerierbar sind. Beispielsweise sind die Elektronen im Atom oder die Neutronen bzw. Protonen im Atomkern identische Teilchen. Der Systemzustand wird dann *vollständig* durch die Gesamtheit aller *Besetzungszahlen* n_k fixiert (↑ 2.5.5.3.). Der Index k numeriert dabei mögliche Zustände und nicht bestimmte Teilchen. Für Boseteilchen oder *Bosonen* sind prinzipiell alle Besetzungszahlen erlaubt, bei Fermiteilchen oder *Fermionen* schränkt das Pauliprinzip die möglichen Werte auf $n = 0$, d. h. unbesetzt, und $n = 1$, d. h. besetzt, ein.

Bei einem *klassischen System* existieren keine diskreten und damit numerierbaren Zustände. Der Zustand eines bewegten Teilchens ohne innere Freiheitsgrade wird durch seinen Ort und seinen Impuls, der eines Systems durch die Orte und Impulse aller es aufbauenden Teilchen bestimmt. Da Ort und Impuls kontinuierlich variable Größen sind, bilden auch die Zustände eines klassischen Systems ein Kontinuum. Daher ist es nicht sinnvoll (↑ 2.4.2.4.), nach der Wahrscheinlichkeit für einen genau bestimmten Zustand zu fragen, sondern nur nach der für einen endlichen, wenn auch beliebig kleinen Zustandsbereich, charakterisiert durch entsprechende Intervalle für Orte und Impulse.

2.4.4.2. Exponentielles Verteilungsgesetz. Im thermodynamischen Gleichgewicht werden alle Mikrozustände des Systems gleicher Energie auch mit gleicher Wahrscheinlichkeit angenommen. Sind beispielsweise die verschiedenen Orientierungen eines magnetischen Dipols (↑ 2.2.2.5.) aufgrund der Abwesenheit eines äußeren Magnetfelds energetisch gleichwertig, so sind sie auch gleichwahrscheinlich. – Kann sich ein Teilchen in jeder von mehreren gleich tiefen Potentialmulden aufhalten, so wird es bei hinreichend langer Meßdauer etwa gleich oft in jeder von ihnen gefunden werden, falls nicht die Barrieren zwischen den Mulden so hoch sind, daß jede Lage bereits für sich einen (im thermodynamischen Sinne) metastabilen Gleichgewichtszustand darstellt. – Fliegt ein Teilchen in einem Kasten mit reflektierenden Wänden, so hält es sich bei langer Beobachtungszeit in gleich großen Teilvolumina des Kastens gleich lange auf (s. a. 2.4.3.1.).

Mikrozustände unterschiedlicher Energie sind nicht gleichwahrscheinlich, vielmehr werden energieärmere Zustände häufiger als solche mit größeren Energien realisiert. Beispielsweise wird sich der Dipol im obigen Beispiel häufiger parallel als antiparallel zu einem angelegten Magnetfeld ausrichten, und in der tiefsten Potentialmulde wird man das Teilchen am häufigsten finden. Solange die Teilchenzahl des Sy-

stems konstant bleibt, wird das Verhältnis w_n/w_m der Wahrscheinlichkeiten zweier Quantenzustände n und m nur von der Differenz $E_n - E_m$ ihrer Energien bestimmt. Die funktionelle Abhängigkeit $w_n/w_m = f(E_n - E_m)$ muß die exponentielle Form $\frac{w_n}{w_m} = e^{-\beta(E_n-E_m)} = \frac{e^{-\beta E_n}}{e^{-\beta E_m}}$ haben, damit sich bei Division von w_n/w_p durch w_m/w_p, also bei Bildung von w_n/w_m, die Energie E_p eines beliebigen anderen Zustands p herauskürzt. In $w_n \sim \exp(-\beta E_n)$ ist β ein Parameter von der Dimension einer reziproken Energie, der positiv sein muß, wenn energiereichere Zustände kleinere Wahrscheinlichkeiten haben. Bei Systemen mit (nach oben) begrenztem Energiespektrum können allerdings sogenannte *Besetzungsinversionen* (↑ 4.1.4.1.) auftreten, die durch negative β-Werte beschrieben werden. Es gilt, daß $\beta = \frac{1}{k_B T}$ mit $k_B = 1{,}380 \cdot 10^{-23}\ \mathrm{JK^{-1}}$ durch die absolute Temperatur des Systems bzw. seiner Umgebung bestimmt ist. k_B wird als *Boltzmannkonstante* bezeichnet.

Bei gegebenem diskretem Energiespektrum $\{E_n\}$ des Systems und bei der Temperatur T ist die normierte Wahrscheinlichkeitsverteilung für ein System im thermischen Gleichgewicht durch die kanonische Verteilung

$$(7)\quad w_n = \frac{1}{Z} e^{-E_n/k_B T} \quad \text{mit} \quad Z = \sum_n e^{-E_n/k_B T}$$

gegeben. Die Größe Z wird als *Zustandssumme* bezeichnet. Sie hängt außer von der Temperatur von allen Parametern ab, die das Energiespektrum beeinflussen, z. B. vom Volumen des Systems oder der Stärke eines angelegten Magnetfelds. Bei einem *klassischen System* gilt das exponentielle Verteilungsgesetz in der analogen Form

$$(8)\quad \Delta w = \frac{\Delta\Gamma}{Z_{kl}} e^{-E/k_B T} \text{ mit } Z_{kl} = \int d\Gamma\, e^{-E/k_B T}$$

für die Wahrscheinlichkeit Δw, das System in einem kleinen Intervall $\Delta\Gamma$ seines Zustandsraums zu finden. E ist die (mittlere) Energie zu diesem Intervall, $d\Gamma$ wird durch das Produkt aller Orts- und Impulskoordinaten bestimmt, die zur Kennzeichnung der klassischen Mikrozustände dienen. Einfache Spezialfälle, die diese allgemeine Vorschrift verdeutlichen, werden später (↑ 2.4.4.6.) behandelt.

2.4.4.3. Energie und Temperatur. Aus $w_n/w_m = \exp[-(E_n - E_m)/kT]$ folgt, daß Zustände etwa gleichwahrscheinlich sind, deren Energien sich nur wesentlich weniger als $k_B T$ unterscheiden. Andererseits ist $w_n \ll w_m$, wenn $E_n - E_m$ größer als einige $k_B T$ ist. Dies kann anschaulich so verstanden werden, daß $k_B T$ die Größenordnung der Energien charakterisiert, die das System bei „thermischen Stößen" mit seiner Umgebung austauscht. Daher sollten die einzelnen Teilchen sich mit kinetischen Energien der Ordnung $k_B T$ bewegen, falls sie bei Stoßprozessen ihre Energie ohne Einschränkungen austauschen können. Denn hätten sie wesentlich größere Energien, so müßten sie diese schließlich durch Stoßprozesse an die Umgebung verlieren; hätten sie wesentlich kleinere, so würden sie „aufgeheizt", bis das thermische Gleichgewicht zwischen System und Umgebung hergestellt ist. Diese Schlußfolgerung wird durch genauere Untersuchungen (↑ 2.4.4.6.) bestätigt. Der Energieaustausch bei Stößen wird bei Fermionen durch das Pauliprinzip behindert, die oben formulierte Bedingung ist also durchaus nicht immer erfüllt.

Die Größe der *thermischen Energie* $k_B T$ ist entscheidend für die im thermischen Gleichgewicht möglichen Prozesse, durch die die qualitativen Merkmale des makroskopischen Systemverhaltens wesentlich bestimmt werden. Bei *Zimmertemperatur* ist $k_B T \approx 0{,}025$ eV, d. h. zum Beispiel, daß Wasserstoffatome sich praktisch immer im Grundzustand befinden, da $k_B T$ klein gegen die zur Anregung oder gar Ionisierung erforderliche Energie von der Größenordnung 10 eV ist. Kompliziertere Atome bzw. Moleküle können entsprechend niedrig liegende Anregungszustände besitzen, die Lichtquanten des beim Übergang vom Anregungs- zum Grundzustand emittierten Lichts haben dann Energien der Ordnung 10^{-2} eV, das entspricht (↑ Abb. 2.2.4-4) Frequenzen im fernen Infrarot: Die von Körpern bei Zimmertemperatur ausgehende Wärmestrahlung ist unsichtbar. Der

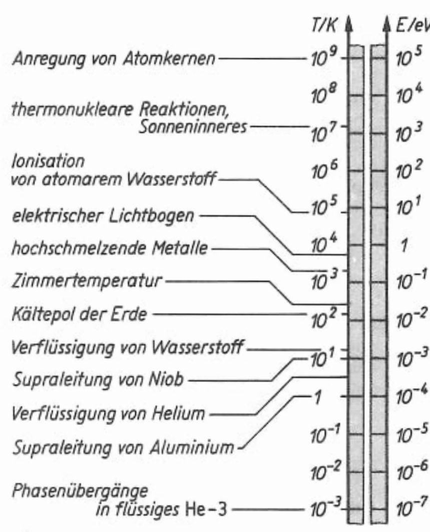


Abb. 2.4.4-1 Größenordnung von Temperatur T und thermischer Energie $E = k_B T$

Übergang zu immer tieferen Temperaturen ist nur möglich und sinnvoll, solange damit neue Ordnungsprozesse in der Materie verbunden sind. Diese Grenze ist möglicherweise im Millikelvinbereich erreicht, in dem sich die Kernspins spontan parallel stellen und neue Phasen von flüssigem He-3 (↑ 6.3.6.1.) gefunden wurden (Abb. 2.4.4-1). Extrem hohe Temperaturen (↑ 9.1.6.) treten im Inneren von Sternen auf, noch weit höhere bestanden vermutlich in früheren Stadien der Entwicklung des Universums.

2.4.4.4. Systeme mit schwankender Teilchenzahl. Falls ein System mit seiner Umgebung sowohl im Energie- als auch im *Teilchenzahlaustausch* steht, so treten die Systemzustände (bei einem System mit nur einer Teilchensorte) im thermischen Gleichgewicht mit Wahrscheinlichkeiten auf, die dem *verallgemeinerten Verteilungsgesetz* $w_{n,N} \sim \exp[\alpha N - \beta E_n^{(N)}]$ genügen. N ist die jetzt schwankende Teilchenzahl, und $E_n^{(N)}$ sind die bei gegebenem N-Wert möglichen Energiewerte. Die Bedeutung von $\beta = (k_B T)^{-1}$ bleibt unverändert. α ist ein *neuer Verteilungsparameter*, der – analog zur Temperatur bzw. zu β beim Energieaustausch – für Teilsysteme, die sich miteinander im thermischen Gleichgewicht befinden, den gleichen Wert hat. α ist also im Sinne der Thermodynamik (↑ 6.1.1.2.) ein intensiver Parameter, dessen Wert man sich durch die Umgebung des Systems, die „Vorratsflasche" für den Teilchenaustausch also, bestimmt denken kann. Mit wachsendem α vergrößert sich die mittlere Teilchenzahl $\bar{N}$ bzw. die zugehörige mittlere Teilchendichte $\bar{N}/V$, wenn V das Volumen des System ist. Umgekehrt kann α als Funktion $\alpha(T, \bar{N}/V)$ von Temperatur und Dichte ausgedrückt werden; es ist in dieser Hinsicht ein bei der theoretischen Behandlung auftretender Parameter, durch dessen Wahl die Teilchendichte im System auf jeden vorgegebenen Wert eingestellt werden kann.

Bei einem makroskopischen System sind die relativen Schwankungen der Teilchenzahl bzw. Dichte meist belanglos (↑ 2.4.2.3.). Daher behandelt man zur Erzielung mathematischer Vorteile oft auch solche Systeme mit dem obigen verallgemeinerten Verteilungsgesetz, deren Teilchenzahl eigentlich fest vorgegeben ist.

2.4.4.5. Zweiniveausystem und harmonischer Oszillator. Die Anwendung der grundlegenden Exponentialverteilung ist dann besonders einfach, wenn das System ein einfaches Energiespektrum besitzt. Gibt es *nur zwei Energieniveaus* mit den Energien E_0 und E_1, so folgt

$$w_0 = \frac{1}{1 + e^{-(E_1 - E_0)/k_B T}}, \quad w_1 = \frac{1}{1 + e^{+(E_1 - E_0)/k_B T}}.$$

Mit steigender Temperatur wächst w_1 und w_0 nimmt ab (Abb. 2.4.4-2). Das System wird durch thermische Stöße aus dem Grundzustand in den angeregten Zustand gehoben. Für $k_B T \gg E_1 - E_0$ ist $w_1 = w_0 = 1/2$: Beide Zustände sind fast gleichwahrscheinlich, da die Differenz ihrer Energien klein gegen die thermische Energie ist. Durch negative absolute Temperaturen werden Situationen mit $w_1 > w_0$ beschrieben.

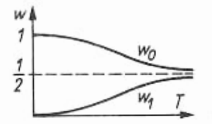


Abb. 2.4.4-2 Wahrscheinlichkeiten für Grund- und angeregten Zustand eines Zweiniveausystems als Funktion der Temperatur

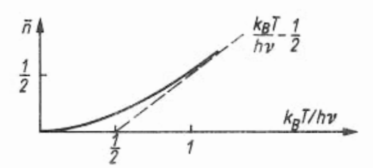


Abb. 2.4.4-3 Mittlere Anregungsenergie eines harmonischen Oszillators als Funktion der Temperatur

Ein harmonischer Oszillator (↑ 2.5.4.3.) hat entsprechend der Planckschen Quantenhypothese (↑ 1.2.5.) ein *äquidistantes Energiespektrum*, bei dem sich benachbarte Niveaus jeweils um die konstante Energiedifferenz $h\nu$ unterscheiden, ν ist die Schwingungsfrequenz des Oszillators. Dementsprechend ist die Wahrscheinlichkeit w_{n+1}, den Oszillator im $(n + 1)$ten Anregungszustand zu finden, jeweils um den Faktor $\exp(-h\nu/k_B T)$ kleiner als die für den n-ten Zustand. Daher ist für niedrige Temperaturen, d. h. für $k_B T \ll h\nu$, bereits w_1 sehr klein gegen $w_0 \approx 1$: Der Oszillator ist praktisch immer im Grundzustand. Dagegen werden die Wahrscheinlichkeiten w_n bei hohen Temperaturen, d. h. bei $k_B T \gg h\nu$, $\exp(-h\nu/k_B T) \approx 1$, mit wachsendem n nur ganz allmählich kleiner: Der Oszillator kann mit vergleichbaren Wahrscheinlichkeiten in sehr vielen, auch hoch angeregten Zuständen gefunden werden.

Der mittlere Anregungsgrad $\bar{n} = \sum_{n=0}^{\infty} n w_n$ kann relativ leicht berechnet werden mit dem Ergebnis

$$(9) \quad \bar{n} = \frac{1}{e^{h\nu/k_B T} - 1} \cong \frac{k_B T}{h} - \frac{1}{2}.$$

Für $k_B T \gg h\nu$ kann die Potenzreihe für die Exponentialfunktion mit dem Glied erster Ordnung abgebrochen werden. So entsteht die angegebene asymptotische Proportionalität von $\bar{n}$ zu $k_B T$ (Abb. 2.4.4-3).

Häufig ist es zweckmäßig, den Übergang in die

nächst höhere Anregungsstufe als *Erzeugung eines Schwingungsquants* zu deuten. n ist dann die Zahl der im entsprechenden Anregungszustand vorhandenen Quanten, $\bar{n}$ deren mittlere Zahl im thermischen Gleichgewicht. Die *Plancksche Formel* [↑ 2.4.(9)] gibt insbesondere die Zahl der Lichtquanten (↑ 1.2.2.) oder auch die der Phononen (↑ 8.1.2.2.) an, die zu einer bestimmten Schwingungsmode, d. h. zu einem bestimmten Wellenzahlvektor (↑ 2.3.1.3.) und zu einem bestimmten Polarisationszustand gehören.

2.4.4.6. Wechselwirkungsfreie klassische Teilchen. Eine Gesamtheit solcher Teilchen bildet ein klassisches ideales Gas. Da jedes einzelne Gasteilchen seine eigene Energie besitzt, kann es als System im Sinne der allgemeinen Begriffsbildungen (↑ 2.4.4.1.) aufgefaßt werden.

2.4.4.6.1. *Räumliche Verteilung.* Befinden sich die Gasteilchen in einem zeitunabhängigen *äußeren Kraftfeld*, das auf alle Teilchen in gleicher Weise wirkt und zu dem das ortsabhängige Potential E_p (↑ 2.1.5.5.) gehört, so ist entsprechend dem exponentiellen Verteilungsgesetz [↑ 2.4.(8)] für jedes Teilchen die Wahrscheinlichkeit, es in einem am Ort $\boldsymbol{r}$ gelegenen Volumenelement ΔV anzutreffen, proportional zu $\Delta V \exp(-E_p(\boldsymbol{r})/k_B T)$. Die gleiche Proportionalität zeigt die mittlere Teilchenzahl in ΔV, so daß die *Dichte der Gasteilchen* sowie der zur Dichte proportionale Druck des inhomogenen idealen Gases entsprechend der Beziehung Druck $\sim$ Dichte $\sim \exp[-E_p(\boldsymbol{r})/k_B T]$ ortsabhängig sind.

2.4.4.6.2. *Impulsverteilung.* Jedes einzelne Gasteilchen ändert bei Stößen mit anderen und mit den Gefäßwänden häufig seinen Impuls. Die Wahrscheinlichkeit Δw_p, daß es bei einer zu beliebiger Zeit vorgenommenen Registrierung einen Impuls besitzt, für den die Spitze des Impulsvektors $\boldsymbol{p}$ in einem kleinen Volumenelement ΔV_{imp} des Impulsraums liegt, ist proportional zu $\exp(-p^2/2mk_B T)$, denn $p^2/2m$ ist die zu dem betrachteten Impulsbereich gehörende kinetische Energie. Zu gleich großen Elementen ΔV_{imp}, deren Impulse sich nur durch die Richtung voneinander unterscheiden, gehören auch gleich große Wahrscheinlichkeiten. Alle Richtungen sind gleichwertig, die *Impulsverteilung ist isotrop.* Faßt man dementsprechend alle Situationen zusammen, bei denen der Impulsbetrag zwischen p und $p + \mathrm{d}p$ liegt, während die Richtung des Impulses beliebig ist, so hat man als Volumenelement im Impulsraum die Kugelschale mit dem Volumen $4\pi p^2\,\mathrm{d}p$ zu wählen. Die Wahrscheinlichkeit, daß der Impuls eines Teilchens der Masse m in dieser Kugelschale liegt, ist also proportional zu $p^2\,\mathrm{d}p \exp(-p^2/2mk_B T)$. Haben alle betrachteten Teilchen die gleiche Masse, so gibt diese Wahrscheinlichkeit gleichzeitig den Anteil $\mathrm{d}N/N$ aller Teilchen an, die sich mit entsprechenden Impulswerten bewegen. Mit dem richtigen, sich aus $\int \mathrm{d}N = N$ ergebenden Normierungsfaktor gilt

$$(10)\quad \frac{\mathrm{d}N}{N} = \frac{4\pi}{(2\pi m k_B T)^{3/2}}\, p^2\, \mathrm{e}^{-\frac{p^2}{2mk_B T}}\, \mathrm{d}p = 4\pi \left(\frac{m}{2\pi k_B T}\right)^{3/2} \cdot v^2\, \mathrm{e}^{-\frac{mv^2}{2k_B T}}\, \mathrm{d}v,$$

wobei in der zweiten Form der *Maxwellschen Geschwindigkeitsverteilung* (↑ 6.2.4.2.) die Teilchengeschwindigkeit v anstelle des Impulses eingeführt worden ist.

2.4.4.6.3. *Gleichverteilungssatz.* Für ein einziges Teilchen stellt 2.4.(10) die normierte Wahrscheinlichkeitsverteilung dar. Für ihren Verlauf ist allein der Quotient $p^2/2mk_B T$ wesentlich. Daher haben schwerere Teilchen wegen des Massenfaktors eine Impulsverteilung, wie sie leichtere Teilchen bei höheren Temperaturen haben, sie besitzen also im Mittel größere Impulse. Dagegen ist die Verteilung der Größe p^2/m bzw. der kinetischen Energie $p^2/2m$ massenunabhängig und allein durch die Temperatur T bestimmt. Insbesondere ist die *mittlere kinetische Energie* für alle Teilchen gleich, und zwar hat sie den Wert

$$(11)\quad E_k = \frac{1}{2m}\,\overline{p^2} = \frac{m}{2}\,\overline{v^2} = \frac{3}{2}\,k_B T.$$

Dieser Gleichverteilungssatz gilt ebenso wie die Maxwellverteilung nicht nur für Gasteilchen, sondern *für jedes klassische Teilchen* in einer beliebigen Substanz oder Umgebung. Auch makroskopische Objekte nehmen an dieser Gleichverteilung teil. Dies führt zu beobachtbaren Schwankungserscheinungen wie der *Brownschen Molekularbewegung* und den thermischen *Schwankungen von Galvanometerspiegeln.*

2.4.4.7. Ideale Quantengase. Auch dann, wenn quantenmechanische Teilchen wechselwirkungsfrei sind bzw. im Rahmen des Modells unabhängiger Teilchen (↑ 2.5.5.1.) so behandelt werden, sind sie dennoch im allgemeinen nicht unterscheidbar, so daß die Resultate des letzten Abschnitts (↑ 2.4.4.6.) nicht anwendbar sind. Die korrekte statistische Fragestellung lautet: Mit welchen Wahrscheinlichkeiten $w_{n(k)}$ nehmen die *Besetzungszahlen* $n(k)$ (↑ 2.4.4.1.) der einzelnen Einteilchenzustände k ihre möglichen Werte an, also $n = 0,1$ bei Fermionen und alle Werte $n = 0, 1, 2 \ldots$ bei Bosonen? Aus den Besetzungszahlen ergeben sich *Teilchenzahl und Gesamtenergie* des Gases durch Summationen $N = \sum_k n_k$, $E = \sum_k \varepsilon_k n_k$ über alle Einteilchenzustände k; entsprechendes gilt für den Zusammenhang der Mittelwerte. Dabei ist ε_k die Energie eines Teilchens im Quantenzustand k.

Alle Teilchen im gleichen Einteilchenzustand können als ein von den restlichen Teilchen aufgrund der Merkmale eben dieses Zustands unterscheidbares Untersystem aufgefaßt werden, das mit dem Rest im Energie- und Teilchenaustausch steht. Daher kann die verallgemeinerte Exponentialverteilung (↑ 2.4.4.4.) auf jedes solches Untersystem angewendet werden, d. h. $w_n \sim e^{\alpha n - \beta \varepsilon n}$. Dabei ist ε eine Eigenschaft des gerade betrachteten Einteilchenzustands, aber α und $\beta = (k_B T)^{-1}$ charakterisieren den Zustand des gesamten Gases. Als wichtigste und für die meisten Fragestellungen bereits ausreichende Größen erweisen sich die mittleren Besetzungszahlen $\bar{n}$, für die sich trotz des formal gleichen Verteilungsgesetzes wegen des unterschiedlichen Wertebereichs der Fermionen- und Bosonen-Besetzungszahlen die verschiedenen Ausdrücke

$$(12) \quad \bar{n} = \frac{1}{\exp\left(\dfrac{\varepsilon - \mu}{k_B T}\right) \pm 1} \quad \text{für} \begin{cases}\text{Bosonen}\\ \text{Fermionen}\end{cases} \quad \text{mit} \quad = \frac{\mu}{k_B T}$$

ergeben. Die Größe μ hat die Dimension einer Energie, sie wird als *chemisches Potential* bezeichnet. Ebenso wie α kann auch μ einfach als ein Parameter zur richtigen Einstellung der mittleren Teilchendichte (↑ 2.4.4.4.) betrachtet werden. Bei Bosonen ist stets $\mu < \varepsilon_{min}$, bei Fermionen besteht eine solche Einschränkung nicht.

Die Gesetzmäßigkeit 2.4.(12) für die Abhängigkeit der mittleren Besetzungszahlen von der Energie wird als *Bose-* bzw. *Fermiverteilung* bezeichnet. Zwar ist deren Verlauf bei gegebenem T und μ übersichtlich (Abb. 2.4.4-4), dennoch sind die physikalischen Aussagen nicht sofort zu übersehen, weil μ keine so einfache physikalische Bedeutung wie die Temperatur T hat. Einfach sind einige Spezialfälle.

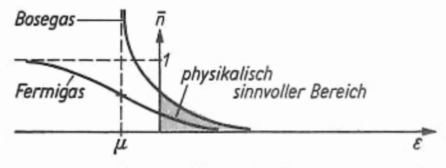


Abb. 2.4.4-4 Fermi- und Boseverteilung für einen gemeinsamen μ-Wert; der kleinste Energiewert wurde als Null angenommen

2.4.4.7.1. *Klassischer Grenzfall.* Falls alle Besetzungszahlen $\bar{n} \ll 1$ klein gegen 1 sind, muß der Exponentialausdruck im Nenner von 2.4.(12) für alle Energien groß gegen 1 sein. Daher kann der Term ± 1 im Nenner vernachlässigt werden, und sowohl die Fermi- als auch die Boseverteilung gehen in die sogenannte *Boltzmannverteilung*

$$(13) \quad \bar{n} \approx e^{(\mu - \varepsilon)/k_B T} \sim e^{-\varepsilon/k_B T}$$

über. Diese rein exponentielle Energieabhängigkeit kennzeichnet ein klassisches ideales Gas (↑ 2.4.4.5.), lediglich in der Darstellung durch die Besetzungszahlen von stationären Einteilchenzuständen sind quantenmechanische Aspekte enthalten. Wenn alle mittleren Besetzungszahlen klein gegen 1 sind, so ist n fast immer Null, selten 1 und fast nie größer als 1, so daß die durch das Pauliprinzip bedingten Unterschiede im Verhalten von Fermionen und Bosonen verschwinden müssen. Tatsächlich werden die Besetzungszahlen dann klein, wenn die Zahl der auf die einzelnen Zustände zu verteilenden Teilchen klein ist gegen die Zahl der Zustände mit vergleichbaren Besetzungswahrscheinlichkeiten. Daher führen kleine Teilchendichten und/oder hohe Temperaturen zu klassischem Verhalten. Für Teilchen mit dem nichtrelativistischen Energie-Impuls-Zusammenhang $\varepsilon = p^2/2m$ muß $\bar{N}/V \ll (mk_B T)^{3/2}/h^3$ sein.

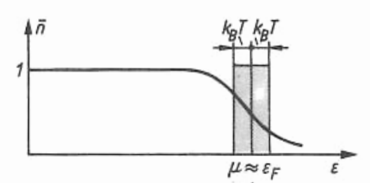


Abb. 2.4.4-5 Verteilungsfunktion eines stark entarteten Fermigases

2.4.4.7.2. *Stark entartetes Fermigas.* Bei der Temperatur $T = 0$ entartet die Fermiverteilung zu einer Stufenfunktion. Alle Zustände mit Energien, die kleiner als die Grenzenergie $\mu(0)$ sind, sind entsprechend $\bar{n} = 1$ mit Sicherheit besetzt, während alle anderen entsprechend $\bar{n} = 0$ mit Sicherheit unbesetzt sind. Das entspricht dem Aufbau des *quantenmechanischen Grundzustands*, der sich für $T \to 0$ herausbildet. Das chemische Potential bei $T = 0$ wird *Fermische Grenzenergie* genannt und meist mit ε_F bezeichnet. Solange bei von Null verschiedenen Temperaturen $k_B T \ll \varepsilon_F$ bleibt, wird die scharfe Grenze zwischen besetzten und unbesetzten Zuständen lediglich über ein Energieintervall der Größenordnung $k_B T$ verschmiert (Abb. 2.4.4-5), während $\mu(T) \approx \varepsilon_F$ näherungsweise noch gleich ε_F bleibt. Die Teilchen in energetisch tiefliegenden Niveaus werden also zunächst thermisch nicht angeregt. Nur solche, deren Energien in einem Intervall der Ordnung $k_B T$ dicht unterhalb der Grenzenergie liegen, können bei Zufuhr einer thermischen Energie $k_B T$ in vorher

unbesetzte und daher dem Pauliprinzip entsprechend aufnahmefähige Zustände „angehoben“ werden.

2.5. Grundideen und Aussagen der Quantentheorie

2.5.1. Quantentheoretisches Schema

2.5.1.1. Welle-Teilchen-Dualismus. Für Licht, Elektronen und andere elementare Objekte existieren unanfechtbare experimentelle Belege (↑ 1.2.) dafür, daß sich diese Objekte einerseits wie Wellen, andererseits wie Teilchen verhalten. Dies findet seinen formelmäßigen Ausdruck in den Beziehungen $E = h\nu$ und $p = h/\lambda$, die erstmals von PLANCK und DE BROGLIE gefunden wurden. Durch diese Relationen werden die Teilchengrößen Energie E und Impuls p mit den Wellengrößen Frequenz ν und Wellenlänge λ verknüpft und das Plancksche Wirkungsquantum h in der Naturbeschreibung verankert. Das Doppelbild der Materie ist ein experimenteller Fakt. Dieser Fakt ist die Wurzel aller Seltsamkeiten der Quantentheorie und gleichzeitig ihre wichtigste experimentelle Grundlage. Sowohl das Teilchen- wie das Wellenbild lassen sich rein klassisch formulieren. Jedes dieser Bilder beschreibt einen Teil der beobachteten Erscheinungen. Die Versuchsbedingungen können so gewählt werden, daß Teilchen- oder Wellencharakter relativ rein hervortreten. Keines der beiden Bilder ist jedoch ausreichend, um allein alle Beobachtungen zu erklären. Erst die gleichzeitige Benutzung beider Bilder für das gleiche Objekt erlaubt, alle Beobachtungen zu interpretieren. Für diese gleichberechtigte Verwendung zweier sich gegenseitig ausschließender anschaulicher Bilder hat BOHR den Begriff Komplementarität geprägt. Der quantenmechanische Formalismus liefert die notwendig unanschauliche Synthese beider Bilder. Die Interpretation dieses Formalismus muß erkennen lassen, worin das Wesen dieser Synthese besteht.

Die Synthese von Teilchen- und Wellenbild stellt ein sehr schwieriges Problem dar, insbesondere, weil Teilchen und Wellen einander ausschließende Strukturen bezeichnen. Ein Objekt kann entweder räumlich scharf lokalisiert, d. h. ein Teilchen, oder räumlich verteilt, d. h. eine Welle, sein. Eine einfache raum-zeitliche Synthese, ein einfaches „Aufeinanderkleben“ beider Bilder ist daher nicht möglich. Andererseits sind alle Versuche erfolglos geblieben, durch einen geeigneten Ausbau mit nur einer der beiden Vorstellungen, der von Teilchen oder der von Wellen, zum Ziel zu gelangen. Die echte Synthese beider Bilder erfordert offenbar eine „neue Dimension“, eine Erweiterung des Begriffssystems der Naturbeschreibung. Den Schlüssel zum erweiterten Begriffsraum liefert wieder das Experiment, das die Existenz von Zufallsprozessen in der Natur zeigt.

2.5.1.2. Zufallsprozesse im atomaren Bereich. Einfache Zufallsprozesse sind das Werfen einer Münze oder das Würfeln. Jede Theorie von Zufallsprozessen beruht auf der Verwendung von *Wahrscheinlichkeiten* (↑ 2.4.1.), die eine quantitative Bewertung von Möglichkeiten liefern.

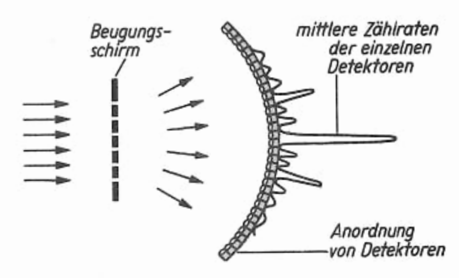


Abb. 2.5.1-1 Registrierung von Einzelteilchen bei einem Beugungsexperiment

Alle Experimente der Mikrophysik beruhen auf dem Nachweis von Mikroobjekten wie Elektronen, Lichtquanten, α-Teilchen durch geeignete Registriereinrichtungen, z. B. Fotoplatten, Geigerzähler, Szintillationszähler. Falls die experimentellen Bedingungen so eingerichtet werden, daß die einzelnen Teilchen getrennt registriert, also gezählt werden, so entstehen in jedem Falle *zufällige Folgen* von Einzelereignissen. So emittiert ein radioaktives Präparat (↑ 3.4.3.1.) die Zerfallsprodukte, bzw. ein Ensemble angeregter Atome bzw. Moleküle die einzelnen Lichtquanten, völlig zufällig. Die radioaktiven Zerfälle bzw. die Lichtquantenemissionen genügen experimentell sehr genau überprüfbaren statistischen Gesetzen, insbesondere zeigen sie ein *exponentielles Zerfallsgesetz* [↑ 2.4.(6)] und eine *Poissonverteilung* [↑ 2.4.(5)] *für die Schwankungen* der Zählrate. Diese Gesetzmäßigkeiten lassen sich mit den Hilfsmitteln der Wahrscheinlichkeitstheorie vollständig aus einfachen statistischen Annahmen ableiten (↑ 2.4.3.4.).

Als zweites Beispiel kann ein beliebiger *Interferenzversuch* dienen (Abb. 2.5.1-1). Sorgt man durch eine entsprechende Vielzahl von Detektoren und eine hinreichend geringe Intensität der Primärstrahlung dafür, daß mit Ausnahme seltener Koinzidenzen immer nur ein Teilchen registriert wird, so sprechen die einzelnen Zähler völlig regellos und unabhängig voneinander an, jedoch so, daß im Mittel über eine sehr große Zahl registrierter Einzelereignisse die Anzahlen der von den einzelnen Detektoren registrierten Teilchen sich zueinander verhalten wie die von der *Materiewellentheorie* vorhergesagten Intensitäten (s. a. 1.2.4.). Der von der Welleninter-

ferenz geprägte *Intensitätsverlauf* wird im statistischen Mittel *durch die Häufigkeitsverteilung* der registrierten Teilchen *realisiert*.

2.5.1.3. Wahrscheinlichkeiten in der Quantentheorie. Da Zufallsprozesse für atomare Erscheinungen typisch sind, muß der Wahrscheinlichkeitsbegriff für ihre Theorie, für die Quantentheorie, wesentlich sein. Tatsächlich beruht die in der Quantentheorie erreichte Synthese von Teilchen- und Wellenaspekten auf dem Wahrscheinlichkeitsbegriff. Die Quantentheorie entwirft kein raum-zeitliches Modell, das etwa das Zustandekommen von für Wellenausbreitungen typischen Interferenzen bei gleichzeitiger Registrierung einzelner Teilchen im üblichen Sinn „erklären" würde. Dieser Verzicht ist grundlegend und typisch für die moderne Quantentheorie. Diese stellt und beantwortet viel mehr die folgenden beiden Fragen:

Welche Werte können bei Messungen einer physikalischen Größe wie Ort, Impuls, Energie, Drehimpuls überhaupt als *Meßwerte* beobachtet werden? Mit welchen *Wahrscheinlichkeiten* treten die einzelnen Meßwerte auf?

Die zweite Frage ist deshalb grundlegend, weil es sich als prinzipiell unmöglich erweist, ein Mikroobjekt so zu „präparieren", daß das Ergebnis einer Messung für alle meßbaren Größen eindeutig bestimmt ist.

Die Antwort auf die erste Frage ist bei einigen Meßgrößen wie Ort und Impuls die gleiche wie bei klassischen Teilchen, es sind wie dort alle Werte prinzipiell möglich. Andere Größen haben nur ein *diskretes Spektrum möglicher Meßwerte*. Das wichtigste Beispiel hierfür sind die *Drehimpulse*. Die *Energie* schließlich ist eine Größe, bei der beide Fälle auftreten: *diskrete Energiewerte* bei räumlich gebundenen Zuständen und ein *kontinuierliches Spektrum* bei freien Zuständen. Die Beobachtung diskreter Energiewerte in den Spektren der Atome gehört zu den ersten überhaupt beobachteten Quanteneffekten; dem diskreten, quantenhaften Charakter der Energieniveaus des harmonischen Oszillators verdankt die Quantentheorie ihren Namen.

Die Wahrscheinlichkeiten für die einzelnen Meßwerte stimmen überein mit den relativen Häufigkeiten in *statistischen Kollektiven aus sehr vielen Meßresultaten*, die alle an gleichartig „präparierten" Mikroobjekten gewonnen wurden, z. B. an Elektronen, die alle die gleichen elektrischen Felder, Blendensysteme, Beugungsgitter u. a. durchlaufen haben.

Zwei quantenmechanische Systeme befinden sich dann im gleichen Zustand, wenn die Wahrscheinlichkeitsverteilungen für alle möglichen Messungen übereinstimmen. Da zur Messung verschiedene Meßanordnungen erforderlich sind, handelt es sich dabei um einen sehr abstrakten Zustandsbegriff.

2.5.1.4. Wahrscheinlichkeitsamplituden. Die quantenmechanischen Wahrscheinlichkeiten werden im Rahmen des quantenmechanischen Formalismus stets so bestimmt, wie es ihrer *Korrespondenz* (↑ 2.4.1., 2.5.1.2.) *zu den Intensitäten* einer Wellenausbreitung entspricht. Die Intensität eines Wellenfelds (↑ 2.3.2.1.) ist im wesentlichen das Quadrat bzw. Absolutquadrat $|u|^2$ der zugehörigen Feldfunktion. Dementsprechend wird die quantenmechanische Wahrscheinlichkeit w_n für das durch den Index n hier nur symbolisch gekennzeichnete Meßergebnis stets als Absolutquadrat $w_n = |c_n|^2$ einer im allgemeinen komplexen *Wahrscheinlichkeitsamplitude* c_n bestimmt. Diese Amplituden genügen linearen Gleichungen und können daher genauso wie die Felder bei einer durch lineare Gleichungen beherrschten Wellenausbreitung überlagert werden. Eine solche *Linearkombination von Amplituden* führt analog zu den klassischen Welleninterferenzen (↑ 2.3.2.1.) zu gemischten Termen in den zugehörigen Wahrscheinlichkeiten, also zu der für die Quantentheorie typischen sogenannten *Interferenz von Wahrscheinlichkeiten*. Dadurch unterscheidet sich die Quantentheorie wesentlich von allen anderen Disziplinen, die ebenfalls mit Wahrscheinlichkeiten arbeiten.

Im mathematischen Formalismus der Quantentheorie treten nur die Amplituden auf. Diese sind selbst nicht direkt meßbar. Dadurch erhalten sie den Charakter von Hilfsgrößen bzw. -funktionen zur Bestimmung der meßbaren Wahrscheinlichkeiten. Offenbar wäre es völlig sinnlos, den zu verschiedenen meßbaren physikalischen Größen gehörenden Wahrscheinlichkeitsverteilungen $\{w_n\}$ *unabhängig* voneinander Amplituden $\{c_n\}$ zuzuordnen, da deren Phasen dann prinzipiell unbestimmbar bleiben müßten. Entscheidend ist vielmehr, daß zwischen den verschiedenen Amplitudensätzen *lineare Relationen* bestehen, wobei schließlich ein einziger Satz von Amplituden zur Berechnung aller anderen Sätze genügt. Die Auswahl dieses einen Satzes erweist sich als weitgehend willkürlich. Ist jedoch ein solcher Satz gefunden, so realisiert er eine vollständige Beschreibung des gerade vorliegenden quantenphysikalischen Zustands (↑ 2.5.1.3.).

Um dieses Schema zu konkretisieren, sind zwei Fragen zu beantworten:

Welche physikalischen Größen können zu einer vollständigen Charakterisierung der bisher durch den Index n nur symbolisch gekennzeichneten Fälle benutzt werden? Dabei verhindert der universelle, prinzipiell auch für die makroskopischen Meßgeräte gültige Welle-Teilchen-Dualismus die gleichzeitige Messung beliebiger physikalischer Größen und damit die beliebige „Verfeinerung" der Beschreibung. Andererseits kann natürlich die Beschreibung des Zu-

stands nicht vollständig sein, wenn bei seiner Fixierung vorhandene Meßmöglichkeiten unberücksichtigt bleiben.
Welche Relationen bestehen zwischen den zu verschiedenen Messungen gehörenden Wahrscheinlichkeitsamplituden?

2.5.1.5. Ortsverteilung und Wellenfunktion. Ein Teilchen kann bei einer Messung prinzipiell an jedem Raumpunkt mit dem Ortsvektor $\boldsymbol{r}$ gefunden werden. Die möglichen Meßwerte des Teilchenorts bilden also ein (dreidimensionales) Kontinuum, genau wie bei einem klassisch beschreibbaren Teilchen. Die zugehörige Wahrscheinlichkeitsamplitude kann daher nicht durch einen diskreten Index n gekennzeichnet werden, sondern muß eine (komplexe) Funktion $\psi(\boldsymbol{r})$ des Orts sein. Diese Funktion heißt *Wellenfunktion* des Teilchens; wie alle Wahrscheinlichkeitsamplituden ändert sie sich im allgemeinen im Zeitablauf. Da es generell, d. h. nicht nur in der Quantentheorie (↑ 2.4.2.4.), bei kontinuierlich variierenden Größen keinen Sinn hat, nach der Wahrscheinlichkeit für einen bestimmten Wert aus diesem Kontinuum zu fragen, ist $\mathrm{d}w = |\psi(\boldsymbol{r})|^2 \, \mathrm{d}V$ die Wahrscheinlichkeit dafür, das Teilchen bei einer Ortsmessung in einem beim Ort $\boldsymbol{r}$ gelegenen Volumenelement $\mathrm{d}V$ zu finden. Die Wellenfunktion wurde zunächst von DE BROGLIE und SCHRÖDINGER zur Beschreibung von Materiewellen (↑ 1.2.3.) eingeführt; ihre *Wahrscheinlichkeitsinterpretation* stammt von BORN. Da das Teilchen bei einer Messung *mit Sicherheit*, d. h. mit der Wahrscheinlichkeit 1, irgendwo gefunden werden muß, gilt das Integral $\int \mathrm{d}w = \int \mathrm{d}V \, |\psi(\boldsymbol{r})|^2 = 1$. Das Integral über $|\psi|^2$ heißt *Normierungsintegral* der Wellenfunktion; die Bedingung, daß dieses Integral gleich 1 ist, heißt *Normierungsbedingung*. Ist der Ort eines Teilchens sehr genau bestimmt, so ist die Wellenfunktion in einem entsprechend kleinen Intervall konzentriert und außerhalb gleich Null. Die Breite des Bereichs, in dem ψ wesentlich von Null verschieden ist, ist ein Maß für die bei Ortsmessungen auftretende Streuung der Meßwerte.

Bei einem einzigen Teilchen entspricht die Wellenfunktion $\psi(\boldsymbol{r})$ weitgehend der Feldfunktion der zugehörigen klassischen Materiewelle. Der Übergang von der klassischen Wellentheorie zur Quantentheorie wird einfach durch die statistische Interpretation der ψ-Funktion vollzogen. Die daraus unmittelbar folgende Normierungsbedingung ist einer reinen Feld- bzw. Wellentheorie völlig fremd. Sie spiegelt den Aspekt wider, daß genau ein Teilchen vorhanden ist.

2.5.1.6. Impulsverteilung und Wellenfunktion. Nach der de-Broglieschen Relation (↑ 1.2.3.1.) $p = \hbar k = h/\lambda$ zwischen Teilchenimpuls p und Wellenzahl k bzw. Wellenlänge λ muß ein Teilchen mit bestimmtem Impuls durch eine Wellenfunktion beschrieben werden, die eine ebene Welle entsprechender Wellenlänge darstellt. Da der Impuls genau so wie der des Teilchens prinzipiell jeden Wert aus einem Kontinuum von Werten annehmen kann, sollte es Wellenfunktionen, die ein Teilchen mit genau fixiertem Impuls beschreiben, eigentlich nicht geben. Dementsprechend ist eine periodische *ebene Welle*, die definitionsgemäß im gesamten unendlichen Raum mit konstanter Amplitude vorhanden ist, nicht normierbar (s. a. 2.5.2.3.). Dagegen kann ein *normierbares Wellenpaket* durch Überlagerung ebener Wellen dargestellt werden. Wird diese Überlagerung gemäß $\psi(x) = (2\pi)^{-1} \int \mathrm{d}k \, \mathrm{f}(k) \, \mathrm{e}^{\mathrm{i}kx}$ vorgenommen (↑ 2.3.1.6.), so ist – im Unterschied zu einer Entwicklung nach Sinus- und Kosinusfunktionen – auch die Richtung des den einzelnen Fourierkomponenten zuzuordnenden Impulses eindeutig fixiert: Zu $\exp(\mathrm{i}kx)$ gehört der Impuls $\hbar k$, zu $\exp(-\mathrm{i}kx)$ der Impuls $-\hbar k$. Diese Festlegungen sind dann richtig, wenn die Zeitabhängigkeit der Welle durch einen Faktor $\exp(-\mathrm{i}\omega t)$ dargestellt wird.

Das allgemeine Schema (↑ 2.5.1.4.) legt zwingend nahe, die *Wahrscheinlichkeit für Impulsmessungen* mit dem Absolutquadrat $|f(k)|^2$ der Fouriertransformierten zu verbinden. Analog zur Ortsmessung, sollte $\mathrm{d}w = \mathrm{const} \, |\mathrm{f}(k)|^2 \, \mathrm{d}k$ die Wahrscheinlichkeit dafür sein, einen Impuls innerhalb des Intervalls $\mathrm{d}p = \hbar \, \mathrm{d}k$ (eindimensional) bzw. des Volumenelements $\mathrm{d}V^{\mathrm{imp}} = \hbar^3 \, \mathrm{d}^3k$ im Impulsraum (dreidimensional) zu finden. Tatsächlich folgt aus den mathematischen Eigenschaften der Fouriertransformation die Identität $\int \mathrm{d}x \, |\psi(x)|^2 = (2\pi)^{-1} \int \mathrm{d}k \, |\mathrm{f}(k)|^2$. Die Impulsverteilung ist richtig normiert, wenn der konstante Faktor in dem obigen Ausdruck für $\mathrm{d}w$ gleich $(2\pi)^{-1}$ wird. Im räumlichen Fall ist k durch den Wellenzahlvektor $\boldsymbol{k}$ zu ersetzen; $\hbar\boldsymbol{k}$ ist der Impulsvektor, $\exp(\mathrm{i}\boldsymbol{kr})$ eine ebene Welle in der durch $\boldsymbol{k}$ bestimmten Richtung. Bei der Integration tritt dann $\mathrm{d}^3k/(2\pi)^3$ an die Stelle von $\mathrm{d}k/2\pi$, $\mathrm{d}^3k = \mathrm{d}k_x \, \mathrm{d}k_y \, \mathrm{d}k_z$ ist das Volumenelement im drei dimensionalen k-Raum.

Auf diese Weise kann aus der komplexen Wellenfunktion $\psi(\boldsymbol{r})$ nicht nur die zu $|\psi(\boldsymbol{r})|^2$ proportionale Ortsverteilung, sondern auch die zu $|\mathrm{f}(\boldsymbol{k})|^2$ proportionale Impulsverteilung bestimmt werden. Beim Übergang von $\psi(\boldsymbol{r})$ zu $\mathrm{f}(\boldsymbol{k})$ ist die für $|\psi(\boldsymbol{r})|^2$ belanglose Phase der komplexen Wellenfunktion wesentlich. Wegen der Umkehrbarkeit der Fouriertransformation ist in $\mathrm{f}(\boldsymbol{k})$ (aber nicht in $|\mathrm{f}(\boldsymbol{k})|^2$) die gleiche Information über den quantenmechanischen Zustand wie in $\psi(\boldsymbol{r})$ enthalten. Meist ist die Darstellung durch $\psi(\boldsymbol{r})$ am besten geeignet.

Im Zusammenhang mit der Darstellung von $\psi(\boldsymbol{r})$ durch Superposition ebener Wellen wurde die *Wellenfunktion* als *eindeutig* angenommen.

Diese Bedingung bleibt auch im weiteren bestehen.
Auf die beiden allgemeinen Fragen zur Beschreibung quantenmechanischer Zustände (↑ 2.5.1.4.) kann nun für *ein* spinloses Teilchen eine Antwort gegeben werden:
Ein Zustand kann entweder durch die Amplitudenfunktion $\psi(\boldsymbol{r})$ für Ortsmessungen oder durch die Amplitudenfunktion $f(\boldsymbol{k})$ für Impulsmessungen charakterisiert werden.
Beide Amplitudenfunktionen sind durch eine Fouriertransformation linear miteinander verknüpft.
Bei *Teilchen mit Spin* muß zum Ort bzw. zum Impuls die Spineinstellung hinzugefügt werden.

2.5.1.7. Unbestimmtheitsrelation. Die Bereiche Δx bzw. Δk, in denen Wellenfunktion $\psi(x)$ bzw. ihre Fouriertransformante $f(k)$ wesentlich von Null verschiedene Werte annehmen, hängen über $\Delta x \Delta k \approx 1$ miteinander zusammen (↑ 2.3.1.6.). Mit der de-Broglieschen Beziehung $p = \hbar k$ folgt $\Delta p = \hbar \Delta k$ und daraus die *Heisenbergsche Unbestimmtheitsrelation* $\Delta p \Delta x \approx \hbar$ für die Schwankungsbreiten von Ort und Impuls. Bei Ortsmessungen streuen also die Meßresultate in einem Intervall Δx, bei Impulsmessungen in Δp. Bei einer starken *Lokalisierung des Teilchens* treten zwangsläufig große Impulse mit nennenswerten Wahrscheinlichkeiten auf. Bei einer ziemlich genauen Festlegung des Impulses erstreckt sich die räumliche Verteilung über einen großen Bereich. Die Unbestimmtheitsrelation gilt bei dreidimensionalen Problemen für die Schwankungsprodukte $\Delta p_x \Delta x$, $\Delta p_y \Delta y$ und $\Delta p_z \Delta z$, während für $\Delta p_x \Delta y$ usw. keine Einschränkungen gelten.
Die Unbestimmtheitsrelation drückt aus, daß es unmöglich ist, quantenmechanische Zustände zu konstruieren, in denen die Ergebnisse von Orts- und Impulsmessungen gleichzeitig beliebig genau vorherbestimmt sind. Der *klassische Bahnbegriff* eines Teilchens ist also *nicht anwendbar*, da ein Teilchen auf einer bestimmten Bahn automatisch zu jeder Zeit einen genauen Ort und einen genauen Impuls hat. Wäre es möglich, ein Teilchen, und sei es nur in einem Gedankenexperiment, in einen Zustand zu versetzen, in dem sein Ort *und* sein Impuls genau bekannt sind, so wäre damit die Inkonsistenz der Quantentheorie erwiesen. Tatsächlich waren alle solchen Versuche erfolglos. Stets konnte man zeigen, daß der durchgängige Teilchen-Welle-Dualismus den Aufbau entsprechender experimenteller Anordnungen verhindert.
Definiert man $\Delta x \equiv \sqrt{\overline{(\Delta x)^2}}$ und $\Delta p \equiv \sqrt{\overline{(\Delta p)^2}}$ präzis als Wurzeln der mittleren Schwankungsquadrate (↑ 2.4.2.3.), die bei Kenntnis der Wahrscheinlichkeitsverteilungen für x und p prinzipiell berechenbar sind, so kann die Heisenbergsche Unbestimmtheitsrelation in der Form der Ungleichung

(1) $\quad \Delta p \Delta x \geqq \hbar/2$

genau gefaßt werden.

2.5.2. Operatoren, Meßwerte und Mittelwerte

2.5.2.1. Impulsoperator. Bei Vorgabe einer beliebigen Wellenfunktion $\psi(\boldsymbol{r})$ haben im allgemeinen weder der Ort noch der Impuls noch irgendeine andere meßbare Größe einen scharf definierten Wert. Bereits bekannt ist (↑ 2.5.1.5. und 2.5.1.6.), wie Wellenfunktionen zu einem nahezu scharfen Orts- bzw. Impulswert beschaffen sind. Kann man systematisch Wellenfunktionen konstruieren, die Zustände mit festen Werten anderer physikalischer Größen beschreiben?
Die Lösung dieses Problems gelingt durch die Einführung des *Impulsoperators* $\hat{p} = (\hbar/\mathrm{i})\,\mathrm{d}/\mathrm{d}x$ bzw. für den dreidimensionalen Fall

$$\hat{p}_x \equiv \left(\frac{\hbar}{\mathrm{i}}\right)\frac{\partial}{\partial x}$$

$$\hat{p}_y \equiv \left(\frac{\hbar}{\mathrm{i}}\right)\frac{\partial}{\partial y} \qquad \hat{\boldsymbol{p}} \equiv \frac{\hbar}{\mathrm{i}}\,\mathrm{grad}\,.$$

$$\hat{p}_z \equiv \left(\frac{\hbar}{\mathrm{i}}\right)\frac{\partial}{\partial z}$$

Dieser Operator $\hat{p}$ hat die Eigenschaft, bei Anwendung auf eine ebene Welle $\exp(ikx)$ mit beliebiger Wellenzahl den zugehörigen Impulswert $\hbar k$ als Faktor zu liefern; bei Anwendung auf die Fourierdarstellung einer beliebigen Wellenfunktion $\psi(x)$ (↑ 2.5.1.6.) wird jede einzelne Fourierkomponente unter dem Integral mit „ihrem" Impulswert multipliziert. Bei einer oszillierenden *Wellenfunktion mit veränderlicher Wellenlänge* und einer Amplitude $A(x)$, die sich erst über Entfernungen von vielen Wellenlängen wesentlich ändert (Abb. 2.5.2-1), liefert die Anwendung von $\hat{p}$ genähert den „lokalen" Impulswert $p(x) = \hbar k(x) = \hbar/\lambda(x)$.

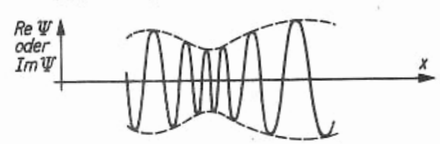


Abb. 2.5.2-1 Wellenfunktion mit schwach ortsabhängiger Amplitude

2.5.2.2. Energieoperator und Schrödingergleichung. Die Energie $E = p^2/2m + V(x)$ des klassischen Teilchens (↑ 2.1.5.5.) setzt sich aus seiner kinetischen Energie $p^2/2m$ und seiner potentiellen Energie $V(x)$ additiv zusammen. Ist das *Potential stückweise konstant*, so gilt

dies auch für das Quadrat des Impulses, und die Beziehung $p = \hbar k$ kann leicht auf jedes Intervall mit konstantem Potential einzeln angewendet werden. Die zugehörige Wellenfunktion genügt daher der Differentialgleichung $\left[\frac{\hat{p}^2}{2m} + V(x)\right] \varphi(x) = \left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right] \varphi(x) = E\varphi(x)$, und zwar längs der ganzen x-Achse, da $\hat{p}$ in jedem Intervall auf die nach Energiesatz und de-Brogliescher Relation zu fordernde Wellenlänge führt. Eine ähnliche Argumentation führt zu dem Schluß, daß diese Differentialgleichung bei einem schwach ortsabhängigen Potential zumindest näherungsweise gültig ist. Als allgemeine Vorschrift zur Ermittlung der *Wellenfunktionen zu bestimmten Energiewerten* bewährt sich diese Gleichung uneingeschränkt. Sie wird als zeitfreie oder *stationäre* (s. a. 2.5.6.1.) *Schrödingergleichung* bezeichnet. Für den dreidimensionalen Fall lautet sie

$$(2) \quad \left[\frac{\hat{p}^2}{2m} + V(\boldsymbol{r})\right] \varphi(\boldsymbol{r}) = \left[-\frac{\hbar^2}{2m}\Delta + V(\boldsymbol{r})\right] \varphi(\boldsymbol{r}) = E\varphi(\boldsymbol{r}) \quad \text{bzw.}$$

$$\left[\frac{1}{2m}(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2) + V(x, y, z)\right] \varphi(x, y, z) = \left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + V(x, y, z)\right] \varphi(x, y, z) = E\varphi(x, y, z),$$

letzteres bei Benutzung kartesischer Koordinaten. Der Laplaceoperator $\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ kann auch in anderen Koordinatensystemen dargestellt werden.

Definiert man den *Energieoperator*, auch *Hamiltonoperator* genannt (↑ 5.1.2.1.), durch $\hat{H} = \frac{\hat{p}^2}{2m} + V(\boldsymbol{r})$, so erhält die Schrödingergleichung die kurze Form $\hat{H}\varphi = E\varphi$: Die Wellenfunktionen zu einem bestimmten Energiewert E sind dadurch vor allen anderen Wellenfunktionen ausgezeichnet, daß die Anwendung des Energieoperators den zugehörigen Energiewert als Faktor liefert, vergleiche die analogen Feststellungen zum Impulsoperator. Der Energieoperator selbst geht bei einem Teilchen in einem Potentialfeld aus dem klassischen Energieausdruck einfach dadurch hervor, daß der Impuls $\boldsymbol{p}$ in dem klassischen Ausdruck für die kinetische Energie durch den Impulsoperator $\hat{\boldsymbol{p}}$ ersetzt wird.

2.5.2.3. Allgemeines Eigenwertproblem. Das eben für die Energie des Teilchens erläuterte Verfahren kann verallgemeinert werden. Dazu wird jeder physikalischen Größe im quantenmechanischen Formalismus ein *Operator* $\hat{A}$ auf folgende Weise zugeordnet: Die betrachtete Größe A wird im Rahmen der klassischen Mechanik als Funktion $A(\boldsymbol{p}, \boldsymbol{r})$ von Ort und Impuls dargestellt. Danach wird der Impuls durch den Impulsoperator ersetzt, wobei gelegentlich noch Symmetrisierungen von Produkten erforderlich sind. Entsprechend verfährt man bei Systemen aus mehreren Teilchen. Die Wellenfunktion hängt dann von den Ortsvektoren *aller* Teilchen ab, und jedes Teilchen hat „seinen" Impulsoperator. Außer Impuls und Energie ist der *Drehimpuls* $\boldsymbol{l} = \boldsymbol{r} \times \hat{\boldsymbol{p}}$ eines Teilchens eine wichtige physikalische Größe. Das erläuterte Schema führt auf die *Drehimpulsoperatoren* $\hat{l}_x = y\hat{p}_z - z\hat{p}_y$, $\hat{l}_y = z\hat{p}_x - x\hat{p}_z$, $\hat{l}_z = x\hat{p}_y - y\hat{p}_x$ für die drei Komponenten.

Bei gegebenem Operator $\hat{A}$ sind diejenigen Wellenfunktionen ψ, die Zustände mit einem schwankungsfreien Wert a dieser Größe beschreiben, als Lösungen der sogenannten *Eigenwertgleichung* $\hat{A}\psi = a\psi$ zu konstruieren. Durch das Auftreten des Impulsoperators handelt es sich dabei um eine Differentialgleichung für $\psi(\boldsymbol{r})$. Stellte man an die Wellenfunktion keine weiteren Forderungen, so könnte a jeden beliebigen, auch komplexen Wert annehmen. Letzteres ist offenbar sinnlos, denn a muß als Wert einer physikalischen Größe reell sein. Fordert man aber *Endlichkeit, Eindeutigkeit und Normierbarkeit*, d. h. hinreichend rasches Verschwinden im Unendlichen *der Wellenfunktion*, so besitzt die Eigenwertgleichung nur noch Lösungen ψ_n für ganz bestimmte diskrete *Eigenwerte* a_n. Für die Energie-Eigenwertgleichung wird dies noch näher erläutert (↑ 2.5.4.2.). Die Eigenfunktionen ψ_n und Eigenwerte a_n können durch einen ganzzahligen Index $n = 0, 1, 2, \ldots$ oder $n = 1, 2, \ldots$ numeriert werden, wobei dieser Index keine weitergehende Bedeutung zu haben braucht. Die *Gesamtzahl der normierbaren Eigenlösungen* kann endlich sein, z. B. für die Energiewerte in einem Potentialkasten endlicher Tiefe, oder unendlich, z. B. für die Energieniveaus des harmonischen Oszillators (↑ 2.5.4.3.) oder die Eigenwerte einer beliebigen Drehimpulskomponente (↑ 2.5.3.1.). Die Operatoren sind stets so beschaffen, daß die *Eigenwerte reell* sind.

Verzichtet man auf die Forderung nach Normierbarkeit der *Eigenfunktionen* (↑ 2.5.1.5.), behält aber die Bedingung bei, daß ψ *überall endlich* bleibt, so können weitere Lösungen der Eigenwertgleichung auftreten. Die zu solchen Lösungen gehörenden Eigenwerte a bleiben reell, bilden aber nun ein *Kontinuum.* Jeder Wert aus diesem kontinuierlichen Wertebereich, der je nach den Eigenschaften des Operators $\hat{A}$ begrenzt oder unbegrenzt sein kann, ist ein möglicher Eigenwert. Das einfachste *Beispiel für ein kontinuierliches Spektrum* liefert der Impuls. Die *Impuls-Eigenwertgleichung* $\hat{p}\psi = (\hbar/i)(d\psi/dx) = p\psi$ hat eine nichtdivergente Lösung für jeden reellen Impulseigenwert p, nämlich die

ebene Welle $\psi = \exp(\mathrm{i}px/\hbar)$. Entsprechend bilden die Meßwerte des Impulses ein unbegrenztes Kontinuum.

Die Gesamtheit der Eigenwerte nennt man das *Spektrum eines Operators* bzw., hier in der Quantenmechanik, der zugehörigen physikalischen Größe. Im allgemeinen hat eine physikalische Größe sowohl ein diskretes als auch ein kontinuierliches Spektrum. Der wichtigste Fall ist die Energie: diskretes Spektrum für die gebundenen Zustände, kontinuierliches Spektrum für die Streuzustände.

2.5.2.4. Meßwerte und Eigenwerte. Die Eigenfunktionen sind Wellenfunktionen mit bestimmtem Wert a der Größe A. Die Eigenwerte sind daher mögliche Meßwerte. Für die Quantentheorie grundlegend ist das weitergehende Postulat:

Die Eigenwerte sind die einzig möglichen *Meßwerte*, und zwar auch dann, wenn sich das System in einem völlig beliebigen Zustand befindet, also durch eine beliebige Wellenfunktion beschrieben wird.

Das ist die Antwort auf die Frage nach den prinzipiell im Experiment auftretenden Meßwerten (↑ 2.5.1.3.).

Die Wahrscheinlichkeit, bei Messungen der Größe A an einem System in einem beliebigen Zustand ψ gerade den Meßwert a_n zu finden ist $w_n = |\int \mathrm{d}V \psi_n^*(r)\, \psi(r)|^2$. Dabei muß sowohl ψ_n wie ψ eine auf 1 normierte Funktion sein (↑ 2.5.1.5.). Damit ist auch die Frage nach den *Wahrscheinlichkeiten der einzelnen Meßwerte* beantwortet.

Die normierbaren Eigenfunktionen des diskreten Spektrums sind mögliche Wellenfunktionen. Ein solcher Eigenzustand kann also physikalisch realisiert werden. Dagegen ist von vornherein klar (↑ 2.4.2.4.), daß bei einem Kontinuum von Meßwerten die Frage nach der Wahrscheinlichkeit für einen mathematisch scharf definierten Wert nicht sinnvoll ist. Dementsprechend sind die Eigenfunktionen zum kontinuierlichen Spektrum nicht normierbar. Durch Überlagerung von Eigenfunktionen aus einem kleinen Intervall des Eigenwertspektrums können normierbare Zustände aufgebaut werden. Bei Messung erhält man dann nur Werte aus diesem Intervall, die Schwankungen sind zur Intervallbreite proportional. Die Überlagerung von Impulseigenzuständen zu normierbaren Zuständen wurde im Fourierintegral (↑ 2.5.1.6.) bereits angewendet.

2.5.2.5. Eigenwerte als Quantenzahlen. Falls zu einem Eigenwert nur eine einzige Eigenfunktion gehört, kann diese (bei gegebenem Operator) durch Angabe dieses Eigenwerts vollständig charakterisiert werden. Die Eigenwerte übernehmen die Rolle von *Quantenzahlen*. Bei gebundenen eindimensionalen Zuständen mit nur einem Teilchen genügt der Energieeigenwert. Bei komplizierteren Problemen gehören oft mehrere voneinander unabhängige Eigenfunktionen zum gleichen Eigenwert. Diese Situation wird als *Entartung* gekennzeichnet. Der Eigenwert allein genügt dann nicht mehr zur eindeutigen Festlegung des Zustands. Dafür braucht man mehrere Quantenzahlen, die häufig Eigenwerte weiterer physikalischer Größen bezeichnen.

Dadurch entsteht die Frage, ob überhaupt mehrere physikalische Größen in *einem* Zustand scharfe Meßwerte haben können, ob also eine Funktion gleichzeitig Eigenfunktion zu verschiedenen Operatoren sein kann. Die Antwort lautet: *Nur vertauschbare Operatoren haben ein gemeinsames System von Eigenfunktionen.* Zwei Operatoren heißen vertauschbar, wenn es nicht auf die Reihenfolge ankommt, in der sie auf eine Funktion angewendet werden. Stehen mehrere Operatoren vor einer Funktion, so ist stets der am weitesten rechts stehende zuerst anzuwenden, d. h., die geforderten Multiplikationen und Differentiationen sind auszuführen.

2.5.2.6. Quantenmechanische Mittelwerte. Der Mittelwert ist als *mittlerer Meßwert* definiert, genauso wie bei einer Benutzung statistischer Methoden außerhalb der Quantenmechanik (↑ 2.4.2.3.). Die möglichen Meßwerte a_n sind nach der Lösung des Eigenwertproblems bekannt, die zugehörigen Wahrscheinlichkeiten bei bekannter Wellenfunktion ψ berechenbar (↑ 2.5.2.3.). Man kann zeigen, daß der Mittelwert entsprechend

$$(3) \qquad \bar{A} = \sum_n w_n a_n = \int \mathrm{d}V \psi^*(r)\, \hat{A} \psi(r)$$

direkt aus $\hat{A}$ und ψ ohne Kenntnis der Lösungen des Eigenwertproblems berechnet werden kann. Daher können Mittelwerte und ähnlich definierte Größen, insbesondere die mittleren Schwankungsquadrate (↑ 2.4.2.3.) meist wesentlich einfacher als die Wahrscheinlichkeitsverteilungen selbst untersucht werden.

2.5.3. Eigenwerte von Drehimpulsen

2.5.3.1. Eigenwerte einer Drehimpulskomponente. Die zugeordneten Operatoren wurden für ein Teilchen bereits angegeben (↑ 2.5.2.2.). Benutzt man die explizite Form $\hat{p}_x = (\hbar/\mathrm{i})\,\partial/\partial x$ usw. der Impulsoperatoren, so kann man durch Übergang zu entsprechenden Winkelkoordinaten für eine beliebige Drehimpulskomponente die Gültigkeit von $\hat{l}_j = (\hbar/\mathrm{i})\,\partial/\partial\varphi_j$ nachrechnen. φ_j ist der Drehwinkel um die j-Achse, also beispielsweise um die z-Achse bei l_z. Der Abstand r vom Zentrum und der Winkel ϑ_j zwischen Ortsvektor und Drehachse sind bei der Differentiation nach φ_j konstant zu halten. Aus dieser Darstellung der Drehimpulsoperatoren wird deut-

lich, daß es nur auf die *Winkelabhängigkeit der Wellenfunktion* ankommt. Insbesondere folgen aus Eigenwertgleichungen für Drehimpulse immer nur Aussagen über die Winkelabhängigkeit der entsprechenden Eigenfunktionen.

Wählt man eine Komponente, z. B. l_z, aus, so ist die *Eigenwertgleichung* $\hat{l}_z\psi = (\hbar/\mathrm{i})\,\partial\psi/\partial\varphi = \hbar m\psi$ zu lösen. Dabei wurden die gesuchten Eigenwerte von l_z in der Form $\hbar m$ angesetzt; m ist dadurch eine dimensionslose Zahl. Die Differentialgleichung $\partial\psi/\partial\varphi = \mathrm{i}m\psi$ hat die Lösung $\psi \sim \exp(\mathrm{i}m\varphi)$, wobei der Proportionalitätsfaktor noch beliebig von r und ϑ abhängen kann. Die Wellenfunktion muß *eindeutig* sein, sie muß also bei einem vollen Umlauf um die Drehachse in sich übergehen. Also muß $\exp(2\pi\mathrm{i}m) = 1$ sein. Dies ist für $m = 0, \pm 1, \pm 2, \pm 3, \ldots$ erfüllt. Bei Messung einer beliebigen Drehimpulskomponente kann als Ergebnis also immer nur ein positives oder negatives ganzzahliges Vielfaches von $\hbar$ gefunden werden.

Zustände mit definiertem m-Wert haben wegen $|\psi_m|^2 \sim |\exp(\mathrm{i}m\varphi)|^2 = 1$ eine drehsymmetrische Wahrscheinlichkeitsverteilung. Der Mittelwert des Ortsvektors hat also keine Komponente senkrecht zur Drehachse. Diese *Gleichwertigkeit aller Richtungen senkrecht zur Achse* zeigt sich allgemeiner darin, daß die *Mittelwerte beliebiger Vektoren* in Achsenrichtung zeigen. Insbesondere gilt dies für den Drehimpulsvektor selbst (s. u.).

Es erweist sich als unmöglich, Wellenfunktionen zu konstruieren, die gleichzeitig Eigenfunktionen zu zwei oder gar zu allen drei Drehimpulskomponenten sind, die sich also beispielsweise bei Drehungen um die x- bzw. y-Achse wie $\exp(\mathrm{i}m_x\varphi_x)$ bzw. $\exp(\mathrm{i}m_y\varphi_y)$ verhalten. Dabei ist zu beachten, daß sich bei einer Drehung des Ortsvektors um die x-Achse auch φ_y ändert und umgekehrt. Eine triviale Ausnahme stellen Zustände dar, deren Wellenfunktion winkelunabhängig ist. Dann ist überhaupt kein Drehimpuls vorhanden. Abgesehen davon gibt es also keine quantenmechanischen Zustände, bei denen mehr als eine Drehimpulskomponente einen genau bestimmten Wert hat. Gilt dies für *eine* Komponente, so verschwinden die quantenmechanischen Mittelwerte der beiden anderen Komponenten, d. h., daß z. B. die Wahrscheinlichkeiten für die Meßwerte $+\hbar m_x$ und $-\hbar m_x$ für alle m_x-Werte gleich sind.

Der tiefere Grund für die Unmöglichkeit, gemeinsame Eigenfunktionen zu zwei oder drei Drehimpulskomponenten zu konstruieren, ist die *Unvertauschbarkeit der Operatoren* $\hat{l}_x$, $\hat{l}_y$ und $\hat{l}_z$ untereinander (↑ 2.5.2.5.). Bei Anwendung auf beliebige Wellenfunktionen gilt $(\hat{l}_x\hat{l}_y - \hat{l}_y\hat{l}_x)\psi(r) = -\dfrac{\hbar}{\mathrm{i}}\,\hat{l}_z\psi(r)$ sowie zwei weitere Vertauschungsrelationen, die aus der angegebenen durch Änderung der Indexfolge, bei Beibehaltung des x-y-z-x-Zyklus, entstehen.

Nach den bisherigen Ausführungen unterscheiden sich die quantenmechanischen Eigenschaften von Drehimpulsen von den aus der klassischen Mechanik bekannten wesentlich, und zwar in zweierlei Hinsicht:

a) Alle Drehimpulskomponenten haben ein *diskretes Spektrum von Meßwerten*. Dieses besteht aus allen ganzzahligen Vielfachen von $\hbar$, ist also besonders einfach.

b) Es gibt prinzipiell keine Zustände, denen ein nach Betrag und Richtung, also mit allen drei Komponenten, genau festgelegter (und nicht verschwindender) Drehimpulsvektor entspräche.

2.5.3.2. Eigenwerte des Drehimpulsquadrats. In einem *halbklassischen Bild* (Abb. 2.5.3-1) kann man das Verschwinden der Mittelwerte von l_x und l_y bei festem Wert $\hbar m$ für l_z durch eine *Präzessionsbewegung des Drehimpulsvektors* auf einem Kegelmantel um die z-Achse veranschaulichen. Gibt man l_z einen der quantisierten Werte $\hbar m$ und ersetzt die Bildung quantenmechanischer Mittelwerte durch die klassischer zeitlicher Mittelwerte über einen oder mehrere Umläufe des Drehimpulsvektors, so enthält das halbklassische Bild die bisherigen Aussagen.

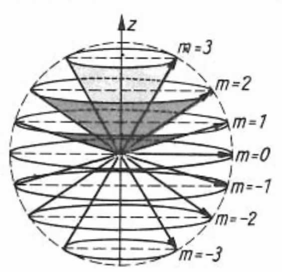


Abb. 2.5.3-1 Halbklassische Darstellung der Drehimpulsquantelung

Mit diesem halbklassischen Bild vereinbar ist eine feste Länge, ein vorgegebener Betrag des Vektors $\boldsymbol{l}$. Daher kann man vermuten, daß auch bei der exakten quantenmechanischen Behandlung $\boldsymbol{l}^2$ und eine Komponente, l_z im folgenden, gemeinsam feste Werte, d. h. Eigenwerte haben können. Dies wird bestätigt: Da der Operator $\hat{\boldsymbol{l}}^2 = \hat{l}_x^2 + \hat{l}_y^2 + \hat{l}_z^2$ mit $\hat{l}_z$, wie natürlich auch mit $\hat{l}_x$ und $\hat{l}_y$, vertauschbar ist, existieren *gemeinsame Eigenfunktionen* von $\hat{\boldsymbol{l}}^2$ und $\hat{l}_z$. Die Eigenwerte von $\boldsymbol{l}^2$ ergeben sich als $\hbar^2 l(l+1)$ mit $l = 0, 1, 2, \ldots$ als natürlicher Zahl. Diese Aussage ist auch dann richtig, wenn man nur das Eigenwertproblem des Operators $\hat{\boldsymbol{l}}^2$, ohne Berücksichtigung von $\hat{l}_z$, löst.

Bei gegebenem Quadrat, also bei fixiertem l-Wert, kann l_z natürlich nicht beliebig groß wer-

den. Es muß $l_z^2 < l^2$ bleiben, für die Quantenzahlen l und m heißt das $m^2 < l(l+1)$. Diese Ungleichung wird für $m = \pm l$ noch erfüllt, für $m = \pm(l+1)$ dagegen verletzt. Die Quantenzahl der z-Komponente durchläuft also die $2l+1$ verschiedenen Werte $m = -l, -l+1, \dots 0, \dots l-1, l$. Dabei ist der Drehimpuls bei gegebenem Betrag verschieden zur z-Achse orientiert. Daher wird m oft als *Orientierungsquantenzahl* bezeichnet. Im halbklassischen Bild (↑ Abb. 2.5.3-1) entstehen Kegel mit verschieden großen Öffnungswinkeln. Auch für die maximalen bzw. minimalen Werte $m = \pm l$ ist dieser Winkel nicht Null; der Drehimpuls kann sich also nicht vollständig in z-Richtung einstellen. Das ist kein Zufall, da Zustände mit betrags- und richtungsmäßig fixiertem Drehimpuls grundsätzlich nicht auftreten können (↑ 2.5.3.1.).

2.5.3.3. Verallgemeinerung, Spins. Die bisherigen Überlegungen gingen von dem Operator für den Bahndrehimpuls eines Teilchens aus. Alle Aussagen über die Eigenwerte bleiben jedoch gültig für Drehimpulse, die durch Vektoraddition beliebig vieler Einzeldrehimpulse aufgebaut werden, sowie für die gar nicht oder nur teilweise auf die Bahndrehimpulse von Bestandteilen zurückführbaren Spins von zusammengesetzten oder elementaren Teilchen (↑ 2.1.3.9.). In all diesen Fällen bleiben die grundlegenden Vertauschungsrelationen zwischen den Operatoren der Drehimpulskomponenten erhalten. Nur aus diesen algebraischen Relationen folgt bereits das Eigenwertspektrum, allerdings mit einer wesentlichen Erweiterung: Außer den ganzzahligen l- und m-Werten können auch *halbzahlige Quantenzahlen*, d. h. $l = 1/2, 3/2, \dots$ auftreten. Während diese Möglichkeit bei Bahndrehimpulsen durch die Forderung nach einer eindeutigen Wellenfunktion ausgeschlossen wird, treten sie bei *Spins* auf. Es ist üblich, bei Spins elementarer bzw. als elementar betrachteter Teilchen die mit dem Drehimpulsquadrat verbundene Quantenzahl mit s anstatt mit l zu bezeichnen, entsprechend m_s bei den Eigenwerten einer Komponente.

Teilchen ohne oder mit ganzzahliger Spinquantenzahl s, kurz mit ganzzahligem Spin, heißen *Bosonen*, solche mit halbzahligem Spin *Fermionen* (↑ 2.5.5.2.). Die für den Aufbau der Materie wesentlichsten Elementarteilchen *Proton, Neutron und Elektron* haben alle den Spin $s = 1/2$ mit den beiden Einstellmöglichkeiten $m_s = \pm 1/2$. *Lichtquanten* haben den Spin 1, aber dem transversalen Charakter der elektromagnetischen Wellen entsprechend ausnahmsweise nur die beiden Einstellmöglichkeiten $m_s = \pm 1$ in Ausbreitungsrichtung. Den Zuständen $m_s = \pm 1$ entsprechen im Wellenbild rechts- bzw. linkszirkular polarisierte Wellen (↑ 2.2.4.2.).

2.5.3.4. Addition von Drehimpulsen. Häufig tritt das Problem auf, zwei Drehimpulse, z. B. die Bahndrehimpulse zweier Teilchen oder den Bahn- und den Spindrehimpuls eines Teilchens, miteinander zu kombinieren. Genau wie in der klassischen Physik addieren sich die z-Komponenten bei der *Bildung der z-Komponente* des resultierenden Drehimpulses. Daher gilt für die zugeordneten Quantenzahlen $M = m_1 + m_2$. Da m_i zwischen $-l_i$ und $+l_i$ $(i = 1, 2)$ variiert, erhält man für M alle Werte zwischen $-(l_1 + l_2)$ und $+(l_1 + l_2)$. Während es für diese extremen M-Werte jeweils nur eine Bildungsmöglichkeit gibt, existieren für die dazwischen liegenden Werte mehrere Möglichkeiten, sie als Summe $m_1 + m_2$ darzustellen: Die Aufteilung der resultierenden z-Komponente auf die beiden Teildrehimpulse ist nicht eindeutig. Die M-Eigenwerte sind dementsprechend im allgemeinen *entartet* (↑ 2.5.2.5.), denn verschiedene Aufteilung bedeutet verschiedene quantenmechanische Zustände. Da jede der insgesamt $(2l_1 + 1)$ Einstellungen von $\boldsymbol{l}_1$ mit jeder der $(2l_2 + 1)$ Einstellungen von $\boldsymbol{l}_2$ kombiniert werden kann, gibt es insgesamt $(2l_1 + 1)(2l_2 + 1)$ Zustände, die sich auf die $2(l_1 + l_2) + 1$ verschiedenen M-Werte verteilen.

Bei der *klassischen Vektoraddition* von $\boldsymbol{l}_1$ und $\boldsymbol{l}_2$ ergibt sich entsprechend der Parallelogrammkonstruktion ein resultierender Vektor $\boldsymbol{L} = \boldsymbol{l}_1 + \boldsymbol{l}_2$, dessen Betrag je nach dem Winkel, den $\boldsymbol{l}_1$ und $\boldsymbol{l}_2$ miteinander einschließen, kontinuierlich zwischen der Summe der Beträge bei genauer Parallelstellung und der Differenz bei genauer Antiparallelstellung beider Vektoren variieren kann. Quantenmechanisch können als Zustände mit bestimmtem Betrag bzw. Quadrat von $\boldsymbol{L}$ nur solche auftreten, bei denen dieses Quadrat einen der möglichen Eigenwerte $\hbar^2 L(L+1)$ annimmt. Der Winkel zwischen $\boldsymbol{l}_1$ und $\boldsymbol{l}_2$ kann also nicht mehr beliebige Werte annehmen, er ist vielmehr durch die Quantenzahlen l_1, l_2 und L bestimmt. Dem halbklassischen Bild entsprechend sollte der maximale bzw. minimale L-Wert den Bedingungen

$$\sqrt{L_{\max}(L_{\max}+1)} \leqq \sqrt{l_1(l_1+1)} + \sqrt{l_2(l_2+1)}$$
$$\sqrt{L_{\min}(L_{\min}+1)} \geqq \sqrt{l_1(l_1+1)} - \sqrt{l_2(l_2+1}$$

genügen. In der zweiten Zeile wurde $l_1 \geqq l_2$ angenommen, was durch entsprechende Numerierung ohne weiteres erreicht werden kann. Die erste Ungleichung wird für $L_{\max} = l_1 + l_2$ noch erfüllt, aber für $L' = l_1 + l_2 + 1$ bereits verletzt. Die zweite wird für $L_{\min} = l_1 - l_2$ erfüllt und für $L'' = l_1 - l_2 - 1$ verletzt. Also sollte L von $|l_1 - l_2|$ bis $(l_1 + l_2)$ variieren.

Diese Vermutung wird durch die strenge quantenmechanische Behandlung des Problems bestätigt. Und zwar nimmt L jeden ganz- bzw. halbzahligen Wert zwischen den angegebenen Grenzen an, das sind $2l_2 + 1$ verschiedene

Werte bei $l_2 \leqq l_1$. Zu jedem L-Wert gibt es $2L + 1$ *mögliche Orientierungen* von L entsprechend den Werten $M = -L \dots + L$ für die z-Komponente. Jeder dieser durch L und M gekennzeichneten Zustände tritt genau einmal auf, d. h., die zugehörige Eigenfunktion kann auf genau eine Weise aus den Eigenfunktionen $\psi_{l_1 m_1}$ und $\psi_{l_2 m_2}$ der Teildrehimpulse aufgebaut werden. Tatsächlich gilt

$$\sum_{|l_1-l_2|}^{l_1+l_2} (2L + 1) = (2l_1 + 1)(2l_2 + 1),$$

die Gesamtzahl aller LM-Zustände stimmt mit der oben diskutierten Zahl der Kombinationsmöglichkeiten von m_1 und m_2 überein.
Diese Behauptung soll am *Beispiel* der Addition zweier Drehimpulse mit $l_1 = 2$ und $l_2 = 1$ erläutert werden. Zu $l_1 = 2$ gehören die 5 Einstellungen $m_1 = -2, -1, 0, +1, +2$, zu $l_2 = 1$ die drei Einstellungen $m_2 = -1, 0, +1$ (Abb. 2.5.3-2). Hinter der geschwungenen Klammer sind die M-Zustände gezeichnet, die man aus allen möglichen Kombinationen von m_1 und m_2 erhält. Beispielsweise bekommt man $M = 2$ zweimal, nämlich aus $\{m_1 = 2, m_1 = 0\}$ und aus $\{m_1 = 1, m_2 = 1\}$. Alle M-Werte zu $L = 3, 2, 1$ treten je einmal auf.

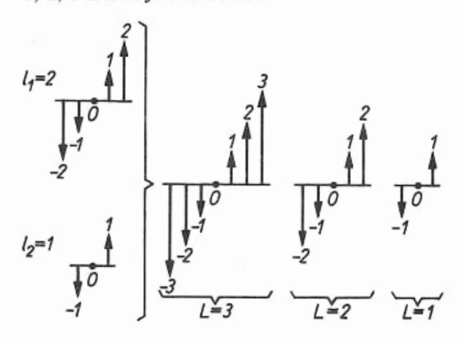


Abb. 2.5.3-2 Kombination der z-Komponenten zweier Drehimpulse

2.5.4. Stationäre Quantenzustände

2.5.4.1. Begriffsbildung. Alle bisherigen Überlegungen betrafen nur die Aussagen, die man aus einer gegebenen Wellenfunktion gewinnen kann, und die Konstruktion besonderer Wellenfunktionen, der Eigenfunktionen. Eine neue und wichtige Frage ist die nach der *zeitlichen Änderung der Wellenfunktion*, also damit auch die nach der zeitlichen Änderung der daraus zu berechnenden Wahrscheinlichkeiten und Mittelwerte. Entsprechend der *Planckschen Beziehung* $E = h\nu = \hbar\omega$ (↑ 1.2.3.1.) liegt es nahe, den Wellenfunktionen zu genau bestimmten Energiewerten E_n eine rein periodische Zeitabhängigkeit mit der Kreisfrequenz $\omega_n = E_n/\hbar$ zuzuordnen. Das geschieht in der Form eines komplexen zeitabhängigen Faktors $\exp(-i\omega_n t)$. Das Minuszeichen ist notwendig, nachdem $\exp(+ikx)$ als der ortsabhängige Anteil einer sich in Richtung der positiven x-Achse ausbreitenden ebenen Welle eingeführt wurde (↑ 2.5.1.6.) Quantenzustände mit genau bestimmter Energie werden beschrieben durch die Eigenfunktionen φ_n des Hamiltonoperators, die der Gleichung $\hat{H}\varphi_n(\boldsymbol{r}) = E_n\varphi_n(\boldsymbol{r})$ genügen (↑ 2.5.2.2.). Befindet sich das Teilchen zur Zeit $t = 0$ im Zustand mit der Wellenfunktion φ_n, so ändert sich diese nach $\varphi_n(\boldsymbol{r}, t) = \varphi_n(\boldsymbol{r})\, e^{-iE_n t/\hbar}$. Die räumliche Wahrscheinlichkeitsverteilung ändert sich durch den Exponentialfaktor nicht, denn es gilt $|\varphi_n(\boldsymbol{r}, t)|^2 = |\varphi_n(\boldsymbol{r})|^2$. Auch die Impulsverteilung und die Verteilungen aller anderen physikalischen Größen bleiben unbeeinflußt. Befindet sich ein System in einem Zustand mit scharfer Energie, so besitzen alle meßbaren Größen zeitunabhängige Wahrscheinlichkeitsverteilungen. In diesem Sinne werden die Eigenlösungen des Energieoperators $\hat{H}$ als *stationäre Zustände*, die Eigenwertgleichung $\hat{H}\varphi_n = E_n\varphi_n$ als *stationäre Schrödingergleichung* bezeichnet. Diese Gleichung hat also eine doppelte Bedeutung: Aus ihr werden das Spektrum der überhaupt möglichen Energiewerte (↑ 2.5.2.2.) und die Wellenfunktionen der stationären Zustände ermittelt.
Die klassische Bewegung des Teilchens erfolgt entsprechend dem Energiesatz mit konstanter, gewissermaßen schwankungsfreier Energie. In dieser Hinsicht entsprechen die stationären Quantenzustände der klassischen Bewegung. Diese Analogie läßt erwarten, daß jede weitere Größe, die bei der klassischen Bewegung neben der Energie auch erhalten bleibt, in einem stationären Quantenzustand ebenfalls einen festen Wert haben kann. Dementsprechend sollten die stationären Zustände Eigenzustände zu dieser weiteren Größe sein. Tatsächlich sind die Operatoren von Erhaltungsgrößen mit dem Hamiltonoperator vertauschbar (↑ 2.5.6.2.), so daß die stationären Zustände außer durch ihren Energiewert noch durch die Werte der *zusätzlichen Erhaltungsgrößen* gekennzeichnet werden können (↑ 2.5.2.5.).

2.5.4.2. Eindimensionale Bindungszustände. Es wurde schon bemerkt, daß die Schrödingergleichung nur für diskrete Energiewerte E_n normierbare Lösungen besitzt (↑ 2.5.2.3.). Eine normierbare Wellenfunktion muß außerhalb eines gewissen Raumbereichs so rasch abklingen, daß das über den unendlichen Raum erstreckte Integral $\int dV\, |\varphi|^2$ endlich bleibt und daher durch geeignete Wahl eines konstanten Faktors bei φ auf den geforderten Wert 1 gebracht werden kann. Die Aufenthaltswahrscheinlichkeit des Teilchens wird also außer-

halb des erwähnten Bereichs ebenfalls sehr klein. Das *Teilchen* kann diesen Bereich praktisch nicht verlassen, es *ist* in diesem Sinne *gebunden*.

Bei der klassischen Bewegung in einem sogenannten *Potentialtopf* (↑ Abb. 2.1.5-3) ist jede Energie oberhalb des Potentialminimums möglich. Für $-V_0 < E < 0$ führt das Teilchen eine Schwingung zwischen den beiden Umkehrpunkten x_1 und x_2 aus (↑ 2.1.5.7.). Der Bereich außerhalb dieses Intervalls ist für ein klassisches Teilchen mit der betrachteten Energie unerreichbar.

Die stationäre Schrödingergleichung des eindimensionalen Potentialtopfproblems ist eine gewöhnliche Differentialgleichung $-\frac{\hbar^2}{2m}\frac{d^2\varphi}{dx^2} \equiv -\frac{\hbar^2}{2m}\varphi'' = [E - V(x)]\,\varphi(x)$.

Außerhalb des Potentialbereichs kann $V \approx 0$ gesetzt werden. So hat für $E < 0$ die Gleichung die *asymptotischen Lösungen* $\varphi \simeq \exp(\pm\varkappa x)$ mit $\varkappa = [2m|E|/\hbar^2]^{1/2}$. Dabei muß man $\varphi \sim \exp(+\varkappa x)$ für $x \to -\infty$ fordern, dagegen $\varphi \sim \exp(-\varkappa x)$ für $x \to +\infty$, weil sonst die φ-Funktion nach außen hin exponentiell anwachsen würde und daher sicher nicht normierbar wäre. Beginnt man links mit $\varphi \sim \exp(+\varkappa x)$, so wird der weitere Verlauf der Wellenfunktion durch die Differentialgleichung vollständig bestimmt. Dabei setzt sich das qualitative Krümmungsverhalten der Exponentialfunktion bis zum linken klassischen Umkehrpunkt x_1 fort. Dort ist $E = V(x)$, also $\varphi'' = 0$, die Funktion hat einen Wendepunkt. Zwischen den Umkehrpunkten, also im klassisch erlaubten Bereich,

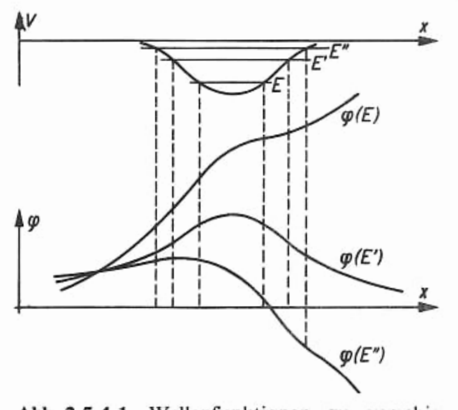


Abb. 2.5.4-1 Wellenfunktionen zu verschiedenen Energiewerten; am linken Schnittpunkt der drei φ-Kurven ist der Wert der Wellenfunktion willkürlich als gleich angenommen worden; $E' = E_0$ Grundzustandsenergie, $E < E_0$ und $E'' > E_0$ liefern keine normierbaren Lösungen

ist $E > V(x)$ und φ'' und φ haben entgegengesetztes Vorzeichen. Die Funktion krümmt sich daher zur x-Achse hin (Abb. 2.5.4-1). Mit wachsender Energie wird die Krümmung stärker, es kommt schließlich zur Ausbildung von Nullstellen und damit zu einem *oszillierenden Verlauf* von $\varphi(x)$. Rechts vom zweiten Umkehrpunkt wird die Funktion wieder exponentialartig, ihr asymptotisches Verhalten ist aber im allgemeinen durch eine Linearkombination $a_+ \exp(+\varkappa x) + a_- \exp(-\varkappa x)$ von abfallender und ansteigender e-Funktion gegeben. Solange $a_+ \neq 0$ ist, bestimmt der zugehörige Term das Verhalten für $x \to +\infty$. Die Wellenfunktion divergiert dann im Unendlichen. a_+ und a_- sind Funktionen von E. Daher bestimmt $a_+(E) = 0$, das ist die *Bedingung für normierbare Lösungen* der Schrödingergleichung, die möglichen diskreten Energiewerte. Für $E > 0$ oszilliert die Wellenfunktion längs der ganzen x-Achse (↑ 2.5.4.7.).

Die *Anzahl der gebundenen Zustände* und ihre Energiewerte hängen von der Form des Potentials ab. Bei dreidimensionalen Problemen kann es bei schwachen Anziehungspotentialen vorkommen, daß es gar keinen gebundenen Zustand gibt. In sehr tiefen und breiten Potentialen existieren entsprechend viele diskrete Energieniveaus. Zum Grundzustand gehört stets eine nullstellenfreie φ-Funktion, zum ersten angeregten Zustand eine Funktion mit einer Nullstelle usw. Der Index n zur Unterscheidung der diskreten Niveaus kann deshalb bei eindimensionalen Problemen so gewählt werden, daß er die Anzahl der Nullstellen der Wellenfunktion angibt.

Die Wellenfunktion und mit ihr auch die Aufenthaltswahrscheinlichkeit $|\varphi|^2$ klingt in dem verbotenen Bereich außerhalb der klassischen Umkehrpunkte exponentiell ab. Dennoch bleibt eine endliche Wahrscheinlichkeit, das Teilchen außerhalb der klassischen Umkehrpunkte anzutreffen. Für schwach gebundene Zustände kann diese Wahrscheinlichkeit sogar wesentlich größer sein als die für den „erlaubten" Bereich. Innerhalb der Umkehrpunkte oszilliert die Wellenfunktion um so rascher, je höher die Energie ist. Die Oszillationen bewirken eine Erhöhung der mittleren kinetischen Energie, da kurzen Wellenlängen große Impulse entsprechen.

2.5.4.3. Potentialkasten und harmonischer Oszillator. Für den Potentialkasten mit unendlich hohen Wänden und das Potential des harmonischen Oszillators können alle Eigenwerte und Lösungen exakt berechnet werden. Da das Potential in beiden Fällen überall positiv ist, sind auch alle Energiewerte positiv. Außerdem geht das Potential für $|x| \to \infty$ gegen unendlich, daher gibt es unendlich viele Niveaus.

Die *Niveaus im Kasten* (Abb. 2.5.4-2) $E_n = \hbar^2\pi^2(n+1)^2/2mL^2$ mit $n = 0, 1, 2, \ldots$ zeigen

eine charakteristische Abhängigkeit von der Kastenbreite L. Die Zustände liegen um so dichter und der Grundzustand um so tiefer, je ausgedehnter das System ist. Die bereits normierten Eigenlösungen $\varphi_{n-1}(x) = \sqrt{2/L}\sin(n\pi x/L)$ sind nur zwischen $x = 0$ und $x = L$ von Null verschieden und verschwinden an den unendlich steilen Potentialwänden. Das Problem ist mathematisch das gleiche wie bei einer an den zwei Punkten $x = 0$ und $x = L$ eingespannten Saite (↑ 2.3.3.1.); die φ_n-Verläufe könnten auch die Grund- und Oberschwingungen einer solchen Saite darstellen.

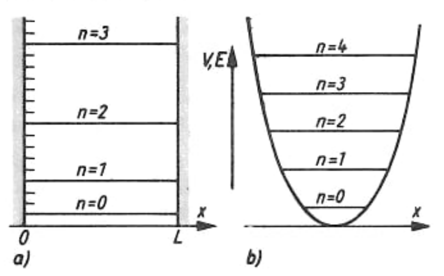


Abb. 2.5.4-2 Energieniveaus; a) für Potentialkasten, b) für harmonischen Oszillator

Beim Oszillator bedeutet $\omega = 2\pi\nu$ die Kreisfrequenz der klassischen Schwingung, das Potential hat die Gestalt $V(x) = m\omega^2x^2/2$, die Kraft deshalb die Form $-m\omega^2x$. Das *Energiespektrum* (↑ Abb. 2.5.4-2) ist $E_n = \hbar\omega(n + \frac{1}{2})$ mit $n = 0, 1, 2, \ldots$

Bei Übergängen zwischen benachbarten Niveaus des Oszillators wird stets die Energie $\Delta E = \hbar\omega = h\nu$ aufgenommen oder abgegeben. In dieser Form wurde die Plancksche Konstante erstmals von PLANCK in die Physik eingeführt, um die Gesetze der Hohlraumstrahlung (↑ 1.2.3.1.) zu erklären. Die Grundzustandsenergie $\hbar\omega/2$ wird häufig als *Nullpunktsenergie* bezeichnet, da sich der Oszillator am absoluten Nullpunkt der Temperatur mit Sicherheit im Grundzustand befindet. Ein klassisches Teilchen würde dann in seiner Gleichgewichtslage $x = 0$ ruhen. Der entsprechende Zustand ist für ein quantenmechanisches Teilchen durch die *Unbestimmtheitsrelation* (↑ 2.5.1.7.) verboten, da dann Ort und Impuls gleichzeitig scharfe Werte, nämlich Null, hätten.

2.5.4.4. Bindungszustände in kugelsymmetrischen Potentialen. Die wesentlichen Merkmale der klassischen Bewegung im Zentralfeld (↑ 2.1.6.3.) werden durch den Energie- und Drehimpulssatz bestimmt. Dem Erhaltungssatz des Drehimpulses entsprechend können die stationären Quantenzustände außer durch ihren Energiewert durch feste Drehimpulswerte charakterisiert werden (↑ 2.5.4.1.). Aufgrund der quantenmechanischen Eigenschaften des Drehimpulses (↑ 2.5.3.2.) kann jedoch nur das Quadrat und eine Komponente feste Werte haben. Mögliche *Eigenwerte des Drehimpulses* sind $\hbar^2l(l+1)$, $l = 0, 1, 2, 3, \ldots$ für das Quadrat und $\hbar m$, $m = -l \ldots +l$ für die Komponente. Da bei einem kugelsymmetrischen Potential alle Raumrichtungen gleichwertig sind, ist die Wahl der wie üblich mit l_z identifizierten Komponente beliebig. Die stationären Zustände können durch die Werte von Energie, Quadrat und z-Komponente des Drehimpulses vollständig charakterisiert werden, wobei die möglichen Drehimpulsquantenzahlen l und m und ihre physikalische Bedeutung von vornherein bekannt sind.

Durchläuft die Orientierungsquantenzahl m bei gegebenem l ihre $2l + 1$ verschiedenen Werte von $m = -l$ bis $m = +l$, so entspricht dies nur einer unterschiedlichen Orientierung des Drehimpulsvektors bezüglich der z-Achse. Bei der klassischen Bewegung hätte das lediglich eine Änderung der Bahnebene zur Folge, ohne weitere Konsequenzen für den Bewegungsablauf in der jeweiligen Bahnebene. Dies macht verständlich, daß auch die Energie der quantenmechanischen Zustände nicht von der Quantenzahl m abhängt (↑ 2.5.4.5.). Es gibt also jeweils $(2l + 1)$ unterschiedliche Drehimpulseinstellungen, also wesentlich verschiedene Wellenfunktionen, zum gleichen Energiewert. Jedes Energieniveau ist (mindestens) $(2l + 1)$-fach entartet.

Der Wert der Betragsquantenzahl l des Drehimpulses beeinflußt die möglichen Energiewerte dagegen wesentlich. Großer Drehimpuls heißt im klassischen Bild rasche Bewegung des Teilchens um das Zentrum, d. h. große kinetische Energie. Auch beim quantenmechanischen Problem wächst die kinetische Energie und damit die Gesamtenergie mit dem Drehimpulsquadrat. Sortiert man daher die Energieniveaus nach den ihnen zugeordneten l-Werten, so gehört der Grundzustand stets zu $l = 0$. Der tiefste Zustand zu $l = 1$ liegt sicher höher als der Grundzustand, er kann aber sowohl über als auch unter dem nächsten Zustand zu $l = 0$ liegen. Entsprechendes gilt für die höheren Drehimpulse. Üblich ist eine von der Spektroskopie herkommende Bezeichnungsweise durch kleine Buchstaben $s, p, d, f, \ldots$, die den Quantenzahlen $l = 0, 1, 2, 3, 4, \ldots$ zugeordnet sind. Die weiteren Buchstaben, folgen in alphabetischer Reihenfolge. Diese Bezeichnungen sind vor der Entwicklung der Quantenmechanik entstanden.

2.5.4.5. Effektivpotential und radiale Schrödingergleichung. Das klassische Bewegungsproblem in einem kugelsymmetrischen Potential (↑ 2.1.6.3.) kann mit Hilfe des Drehimpulssatzes auf ein eindimensionales Problem für die radiale Koordinate r, das ist der Abstand des Teilchens vom Kraftzentrum, zurückgeführt

werden. Die radiale Bewegung erfolgt in einem *Effektivpotential* $V_{\text{eff}}(r) = V(r) + L^2/2mr^2$, wobei das *Zentrifugalpotential* $L^2/2mr^2$ im ursprünglichen Energieausdruck den mit der Winkelbewegung verbundenen Anteil der kinetischen Energie darstellt.

Bei der quantenmechanischen Behandlung gelingt eine analoge Reduktion des zunächst dreidimensionalen Problems, die Lösung $\Phi(\boldsymbol{r})$ der stationären Schrödingergleichung zu bestimmen, auf ein eindimensionales Problem. Der klassischen Trennung in Winkelbewegung und radiale Bewegung im Effektivpotential entspricht die *Produktdarstellung der Wellenfunktion* $\Phi_{Elm}(r, \vartheta, \varphi) = R_{El}(r)\, Y_{lm}(\vartheta, \varphi)$ bei Benutzung sphärischer Kugelkoordinaten r, ϑ und φ. Dabei sind die winkelabhängigen Drehimpulseigenfunktionen Y_{lm} von vornherein bekannt; ihr Verlauf ist hier ohne Interesse. Definiert man durch $R_{El}(r) = r^{-1}u_{El}(r)$ eine neue *radiale Wellenfunktion* $u_{El}(r)$, so genügt diese der *radialen Schrödingergleichung*

$$(4) \quad \left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{\hbar^2 l(l+1)}{2mr^2} + V(r)\right] u_{El} = E u_{El},$$

die tatsächlich einem eindimensionalen Problem für die radiale Koordinate r entspricht. Das beim Zentrifugalpotential als Faktor auftretende Drehimpulsquadrat ist nunmehr entsprechend den möglichen Werten $\hbar^2 l(l+1)$ dieser Größe quantisiert.

Die *Drehimpulsabhängigkeit* des Effektivpotentials wurde bereits für die klassische Bewegung (↑ 2.1.6.3.) diskutiert. Gebundene Zustände sind möglich, wenn das Effektivpotential ein genügend tiefes Minimum durchläuft. Dieses Potentialminimum verschiebt sich mit wachsendem l nach außen und wird flacher, unter Umständen verschwindet es ganz. Letzteres gilt für jeden kugelsymmetrischen Potentialtopf endlicher Breite und Tiefe. Daher existieren in einem solchen Potential gebundene Zustände nicht für beliebig große l-Werte. Das eigentliche Bindungspotential ist dann nicht mehr stark genug, um die bei $l \neq 0$ zwangsläufig auftretenden abstoßenden Zentrifugaleffekte zu kompensieren. Die Aufenthaltswahrscheinlichkeit für Zustände mit größerem Drehimpuls ist bei größeren Abständen vom Zentrum konzentriert.

Die Energieeigenwerte sind aus der radialen Schrödingergleichung [↑ 2.5.(4)] zu bestimmen. Ihre Lösungen $u_{El}(r)$ müssen bei $r = 0$ und für $r \to \infty$ verschwinden. Falls das Potential $V(r)$ für große Abstände verschwindet, verhält sich u wie $u \simeq \text{const} \cdot \exp(-\varkappa r)$, $\varkappa = [2m|E|/\hbar^2]^{1/2}$. Genau wie bei einem gewöhnlichen eindimensionalen Problem (↑ 2.5.4.2.) ist dies nur für ausgewählte *diskrete Energiewerte* möglich. Da die Quantenzahl l selbst explizit in der radialen Schrödingergleichung auftaucht, ist nun auch mathematisch deutlich, daß zu verschiedenen l-Werten im allgemeinen auch verschiedene Energiewerte gehören (↑ 2.5.4.4.).

Falls zu einem gegebenen l-Wert überhaupt gebundene Zustände, d. h. asymptotisch verschwindende Lösungen der radialen Gleichung, auftreten, so hat die zu dem niedrigsten Energiewert gehörende Lösung nur die Nullstelle bei $r = 0$. Die Lösung zur nächst höheren Energie hat eine zusätzliche Nullstelle und so fort. Die Energieniveaus zu festem l können also durch eine radiale Quantenzahl $n_r = 0, 1, 2, \dots$ voneinander unterschieden werden. Diese gibt die Zahl der Kugelflächen mit von Null verschiedenem Radius an, auf denen die Wellenfunktion und damit auch die Aufenthaltswahrscheinlichkeit verschwindet. Die Wellenfunktion und Energieniveaus können entsprechend $\varphi_{n_r lm}$, $u_{n_r l}$, $E_{n_r l}$ indiziert werden.

2.5.4.6. Isotroper harmonischer Oszillator als Beispiel. Ein dreidimensionaler isotroper Oszillator wird durch ein Teilchen im Potential $V(r) = m\omega^2 r^2/2$ realisiert. Dieses Potential ist kugelsymmetrisch; außerdem läßt es sich wegen $r^2 = x^2 + y^2 + z^2$ als einfache Summe von drei eindimensionalen Oszillatorpotentialen (↑ 2.5.4.3.) schreiben. Wegen dieser *Entkopplung* besteht die klassische Bewegung in einer einfachen Überlagerung von drei harmonischen Schwingungen in x-, y- und z-Richtung. Analog dazu ergeben sich die quantenmechanischen Energieeigenwerte einfach als Summe $E = \hbar\omega(n_x + n_y + n_z + 3/2)$ der Energien von drei eindimensionalen Oszillatoren, die unabhängig voneinander angeregt werden.

Der Grundzustand hat die Energie $3\hbar\omega/2$, entsprechend der Nullpunktsenergie $\hbar\omega/2$ für jeden der drei eindimensionalen Oszillatoren. Darüber liegen genau wie im eindimensionalen Fall die

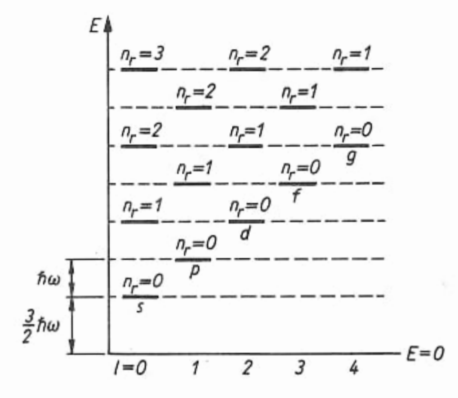


Abb. 2.5.4-3 Lage und Klassifizierung der Energieniveaus des dreidimensionalen isotropen harmonischen Oszillators

angeregten Zustände in Abständen von $\Delta E = \hbar\omega$. Jedes dieser höheren Niveaus ist entartet (↑ 2.5.2.5.), da man verschiedene Wertetripel $\{n_x, n_y, n_z\}$ mit der gleichen Summe finden kann; z. B. gehören $\{2, 0, 0\}$, $\{0, 2, 0\}$, $\{0, 0, 2\}$; $\{1, 1, 0\}$, $\{1, 0, 1\}$ und $\{0, 1, 1\}$ zum zweiten angeregten Zustand. Die Zuordnung von Drehimpulsen, also von l-Werten, zu den einzelnen Niveaus muß, das allgemeine Schema (↑ 2.5.4.4.) konkretisierend, so erfolgen, daß die gleichen Entartungszahlen entstehen. Das Ergebnis kann durch $E_{n_r l} = \hbar\omega(2n_r + l + 3/2)$ dargestellt werden (Abb. 2.5.4-3). Das Zusammenfallen von Niveaus zu verschiedenen l-Werten ist keine Folge der Kugelsymmetrie des Potentials. Diese sogenannte *zufällige Entartung* ist an die spezielle Abstandsabhängigkeit $V \sim r^2$ gebunden und wird durch jede Abweichung von diesem Verlauf aufgehoben.

2.5.4.7. Reflexion und Transmission. Neben Bewegungen, bei denen das Teilchen einen gewissen Raumbereich nicht verlassen kann, existieren klassische Bewegungen, bei denen sich das Teilchen schließlich beliebig weit von jedem Punkt, insbesondere beliebig weit vom Potentialbereich, entfernt. Falls das Potential im Unendlichen verschwindet, sind alle Bewegungen mit positiver Energie von dieser Art. Beispiele sind die Kometenbahnen der Himmelsmechanik und die elastische Streuung von Teilchen an beliebigen Objekten.

Auch bei einem *quantenmechanischen Streuprozeß* kann die Energie der zu streuenden Teilchen beliebig gewählt werden. Bei jedem konkreten Streuexperiment ist sie durch die experimentelle Anordnung, etwa die von Elektronen durchlaufene Beschleunigungsspannung, festgelegt. Dementsprechend müssen quantenmechanische Streuprozesse durch die nicht normierbaren Lösungen der Schrödingergleichung zum kontinuierlichen Energiespektrum (↑ 2.5.2.2.) beschrieben werden. Die Nichtnormierbarkeit der Wellenfunktionen entspricht der Tatsache, daß sich das Teilchen im gesamten Raum aufhalten kann.

Als Beispiel wird die Streuung an einem eindimensionalen Potential $V(x)$ betrachtet, das für $|x| \to \infty$ verschwindet. Für $E > 0$ ist die Situation außerhalb des Potentialbereichs die gleiche wie bei freien Teilchen: Das klassische Teilchen bewegt sich entweder in positiver oder in negativer x-Richtung mit dem Impuls $p = +\sqrt{2mE}$ bzw. $p = -\sqrt{2mE}$. Entsprechend hat die Schrödingergleichung ebene Wellen $\exp(\pm ikx)$ mit $k = p/\hbar = \sqrt{2mE/\hbar^2}$ als *asymptotische Lösungen*. Gibt man auf einer Seite des Potentialbereichs, etwa für $x \to -\infty$, eine beliebige Linearkombination $a_+ \exp(+ikx) + a_- \exp(-ikx)$ vor, so führt die Integration der Schrödingergleichung auf eine dann eindeutig bestimmte Kombination $b_+ \exp(ikx) + b_- \exp(-ikx)$ auf der anderen Seite. Bei einer „von links" einlaufenden Primärwelle $\exp(ikx)$ muß die zu bestimmende Lösung das asymptotische Verhalten

$$\varphi(x) \cong \begin{cases} e^{ikx} + A\,e^{-ikx} & \text{für} \quad x \to -\infty \\ B\,e^{ikx} & \phantom{\text{für}} \quad x \to +\infty \end{cases}$$

haben: „links" Primärwelle und reflektierte Welle, „recht" nur die in gleicher Richtung wie die Primärwelle weiterlaufende durchgelassene Welle. Mit dem Amplitudenfaktor 1 bei der Primärwelle ist in zweckmäßiger Weise über einen ohnehin unbestimmten konstanten Faktor der Wellenfunktion verfügt worden. Das die obige Lösung von anderen mathematisch möglichen Lösungen unterscheidende Merkmal ist das Verschwinden des Koeffizienten b_-. Löst man die stationäre Schrödingergleichung längs der gesamten x-Achse, so erhält man genau bestimmte, im allgemeinen komplexe Werte für die Koeffizienten A und B der reflektierten bzw. durchgelassenen Welle. Beide hängen bei gegebenem Streupotential von der Energie E ab.

Die diskutierte nichtnormierbare Lösung $\varphi(x)$ beschreibt einen stationären Streuvorgang, bei dem ständig ein Strom einfällt, reflektiert und gestreut wird, so wie es praktisch bei Streuexperimenten auch der Fall ist. Der Strom in einer Welle ist proportional zum Amplitudenquadrat $|\varphi|^2$, das der klassischen Teilchendichte entspricht, und zur Geschwindigkeit der Teilchen. An Stelle der Geschwindigkeit kann man den Impuls oder die Wellenzahl k verwenden, denn wegen $p = mv = \hbar k$ unterscheiden sich diese Größen nur um konstante Faktoren. Zur oben angegebenen φ-Funktion gehören danach die folgenden *Ströme* $I_{ein} = 1 \cdot k$, $I_{refl} = |A|^2 k$, $I_{durch} = |B|^2 k$. Man bezeichnet $R \equiv I_{refl}/I_{ein}$ als *Reflexionskoeffizient* und $D \equiv I_{durch}/I_{ein}$ als *Durchlaß-* oder *Transmissionskoeffizient*. Der einfallende Strom teilt sich in reflektierten und durchgelassenen, so daß $I_{ein} = I_{refl} + I_{durch}$ gilt oder $R + D = 1$. Während ein klassisches Teilchen je nach seiner Energie entweder reflektiert oder durchgelassen wird, besteht quantenmechanisch für beide Prozesse eine Wahrscheinlichkeit, nämlich R bzw. D.

2.5.4.8. Tunneleffekt. Der Durchlaßkoeffizient ist auch dann nicht Null, wenn das Potential eine abstoßende Barriere bildet, die das klassische Teilchen nicht durchdringen kann (Abb. 2.5.4-4). Bei Anwendung klassischer Vorstellungen scheint es, als ob sich das Teilchen einen „Tunnel" durch den Potentialberg grübe. Die *Durchlässigkeit solcher Potentialschwellen* wird daher als Tunneleffekt bezeichnet. Aus dem Verlauf von Potential und Wellenfunktion beim Tunneleffekt (in Abb. 2.5.4-4 für eine von links einfallende Welle dargestellt) ersieht man, daß

die Durchlaßwahrscheinlichkeit D um so kleiner wird, je länger der Tunnel und je höher der Potentialberg über dem Tunnel ist. Falls D schon sehr klein ist, kann die Formel

$$D = \exp\left\{-\frac{2}{\hbar}\int\limits_a^b \sqrt{2m[V(x) - E]}\,\mathrm{d}x\right\}$$

zur Abschätzung dienen. Danach gibt in grober Näherung $D \approx \exp\{-(\Delta L/\hbar)\sqrt{m\,\Delta E}\}$ die Durchlaßwahrscheinlichkeit in Abhängigkeit von der „Tunnellänge" ΔL, von der „fehlenden" Energie ΔE und von der Teilchenmasse m wieder. Wachsen diese Größen an, so wird D, wie die Formel zeigt, rasch unmeßbar klein. Das Auftreten der Planckschen Konstante $\hbar$ im Nenner des Exponenten kennzeichnet den Tunneleffekt als typischen *Quanteneffekt*. Solange D und R von gleicher Größenordnung sind, beobachtet man eine nicht monotone Energieabhängigkeit.

Der Tunneleffekt ist für viele Prozesse von großer Bedeutung. Zum Beispiel müssen geladene Teilchen eine durch die elektrostatische Abstoßung hervorgerufene *Coulombbarriere* überwinden, um in einen Atomkern eindringen zu können und dort eine Kernreaktion auszulösen. Umgekehrt muß ein α-Teilchen, das einen radioaktiven Kern beim α-Zerfall verläßt, diesen Wall von innen nach außen überwinden.

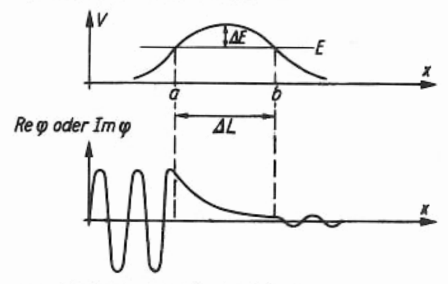


Abb. 2.5.4-4 Zum Tunneleffekt

2.5.5. Identische Teilchen

2.5.5.1. Modell unabhängiger Teilchen. Fast alle wirklich auftretenden Systeme wie Kerne, Atome, Moleküle, Festkörper bestehen aus mehreren bzw. aus sehr vielen Teilchen. Als Beispiel diene die Atomhülle: An den Z-fach positiv geladenen Kern sind Z Elektronen gebunden; Z ist die Kernladungs- oder Ordnungszahl. Die Gesamtheit dieser Elektronen ist eigentlich durch *eine* Wellenfunktion zu beschreiben, die von den Koordinaten aller Elektronen in komplizierter Weise abhängt. Die Bestimmung der exakten stationären Zustände eines solchen *Vielelektronensystems* ist ein praktisch unlösbares Problem Die entscheidende Komplikation wird durch die Wechselwirkung der Elektronen untereinander hervorgerufen. Diese bewirkt, daß die Aufenthaltswahrscheinlichkeit für ein Elektron vom Ort aller anderen Elektronen abhängt.

Diese Schwierigkeit wird durch das Modell der unabhängigen Teilchen umgangen: Die Wechselwirkung eines Elektrons mit allen anderen wird dabei nicht exakt berücksichtigt, sondern durch ein *mittleres Kraftfeld* ersetzt. In diesem Modell finden keine „Stöße" zwischen den Elektronen mehr statt, vielmehr bewegt sich jedes Elektron in einem Kraftfeld, das vom Kern und der mittleren Ladungswolke aller anderen Elektronen hervorgerufen wird. Diese befinden sich ebenfalls in dem mittleren Potential; ihre räumliche Verteilung und damit die Gestalt der ganzen Ladungswolke wird durch das mittlere Potential bestimmt, das andererseits vom Kern und eben dieser Wolke erzeugt wird. Man hat daher das mittlere Potential so lange zu verändern, bis die in ihm sich ausbildende Ladungswolke das gleiche Potential auch wieder erzeugt. Dann erst ist das Verfahren konsistent. Es wird als Methode des *selbstkonsistenten Felds* bezeichnet und ist nicht nur bei der Behandlung der Atomhülle anwendbar.

2.5.5.2. Pauliprinzip. Sind die möglichen Einteilchenzustände im mittleren Potential bestimmt, so ist damit ein Satz von Wellenfunktionen $\varphi_0, \varphi_1, \varphi_2, \ldots$ und zugehörigen Energiewerten $\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots$ gegeben. Die zu verschiedenen Funktionen φ_k gehörenden Energien ε_k können teilweise gleich sein; z. B. treten solche Entartungen zwangsläufig bei allen kugelsymmetrischen Problemen auf (↑ 2.5.4.4.). Wäre der Grundzustand des Systems derjenige Zustand, in dem sich alle Teilchen im untersten Einteilchenzustand φ_0 mit der Energie ε_0 befinden, so wäre die Bindungsenergie eines Atoms etwa zu Z^3, die Ionisationsenergie zu Z^2 und die Atomradien zu Z^{-1} proportional, im Widerspruch zur Erfahrung. PAULI erkannte, daß man die Gesetzmäßigkeiten des Atombaus mit dem folgenden Prinzip verstehen kann:

Jedes nicht entartete, durch eine Wellenfunktion φ_k gekennzeichnete Niveau darf mit maximal zwei Elektronen besetzt werden.

Das ist das Pauliprinzip in seiner ursprünglichen Formulierung. Im Grundzustand werden die Energieniveaus der Reihe nach aufgefüllt, bis alle Elektronen ihren Platz bekommen haben.

Offenbar wäre es ohne eine weitere Erklärung sehr merkwürdig, daß sich gerade *zwei* Elektronen im gleichen Zustand befinden können. Tatsächlich liefert der *Elektronenspin* das fehlende Unterscheidungsmerkmal für zwei solche Elektronen: Sie unterscheiden sich durch ihre Spineinstellung, d. h. durch die Orientierungsquantenzahl $m_s = \pm 1/2$ (↑ 2.5.4.4.).

Das Pauliprinzip gilt in der Form „Ein durch eine ortsabhängige Wellenfunktion φ_k und eine Spineinstellung m_s gekennzeichneter Einteilchenzustand kann nur durch 1 Teilchen besetzt werden" für alle Teilchen, elementare oder zusammengesetzte, mit halbzahliger Spinquantenzahl s, z. B. auch für Atomkerne. Elektronen und Nukleonen haben $s = 1/2$. Allgemein kann für Teilchen mit halbzahligem Spin, d. h. $s = 1/2, 3/2, \ldots$ jeder räumliche Zustand φ_k maximal $(2s + 1)$-fach besetzt werden. Teilchen mit halbzahligem Spin heißen *Fermionen*.

Für Teilchen ohne oder mit ganzzahligem Spin, d. h. $s = 0, 1, 2, \ldots$, gilt das Pauliprinzip nicht, es kann also jeder Einteilchenzustand beliebig oft besetzt werden. Solche Teilchen heißen *Bosonen*. Im Grundzustand eines Bosesystems nehmen alle Teilchen den untersten Einteilchenzustand ein.

Die Struktur des Grundzustands und der unteren angeregten Zustände eines aus identischen Teilchen aufgebauten Systems wird entscheidend durch die vom Teilchenspin abhängigen Besetzungsmöglichkeiten der Einteilchenzustände geprägt. Auf diese Weise entsteht eine *Spinabhängigkeit*, die nichts mit einer etwaigen Spinabhängigkeit der Wechselwirkungskräfte zwischen den Teilchen zu tun hat. Auch bei Systemen im thermodynamischen Gleichgewicht wirken sich die unterschiedlichen Eigenschaften von Fermionen und Bosonen bei tiefen Temperaturen stark aus, da dann nur niedrig angeregte Zustände mit nennenswerten Wahrscheinlichkeiten auftreten (↑ 2.4.4.6.).

2.5.5.3. Besetzungszahlen. Der quantenmechanische Zustand eines aus unterscheidbaren Teilchen bestehenden Systems ist durch eine Wellenfunktion $\Phi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots)$ zu beschreiben. Dabei haben die Orte $\boldsymbol{r}_n$ wirklich die Bedeutung der Orte der entsprechenden numerierten Teilchen. $|\Phi|^2 \, \mathrm{d}V_1 \, \mathrm{d}V_2 \ldots$ ist die Wahrscheinlichkeit, gleichzeitig Teilchen 1 im Volumenelement $\mathrm{d}V_1$ am Ort $\boldsymbol{r}_1$ und Teilchen 2 in $\mathrm{d}V_2$ bei $\boldsymbol{r}_2$ usw. anzutreffen. Bei identischen Teilchen ist eine solche Numerierung nicht mehr möglich. Im Rahmen der klassischen Mechanik wäre die *Identität der Teilchen* noch kein Hindernis für ihre Numerierung; z. B. könnten die Orte und Geschwindigkeiten der Teilchen zu irgendeiner Zeit als Unterscheidungsmerkmal benutzt werden. Da jedes Teilchen dann eine bestimmte Bahn durchläuft, könnte es zu einer späteren Zeit im Prinzip wieder identifiziert werden. In der Quantenmechanik ist der Bahnbegriff unanwendbar und diese Möglichkeit besteht daher nicht. *Identische quantenmechanische Teilchen verlieren ihre Individualität* völlig. Eine konsequente quantenmechanische Beschreibung muß daher von vornherein auf die Einführung von Teilchenindizes verzichten.

Numerierbar und damit durch einen Index k indizierbar bleiben – zunächst im Rahmen des Modells unabhängiger Teilchen (↑ 2.5.5.1.) – die Einteilchenzustände. Es ist sinnlos, den Systemzustand dadurch zu kennzeichnen, *welche* Teilchen die verschiedenen Einteilchenzustände besetzen, sondern nur, *wieviele* Teilchen es jeweils sind. Die *Teilchen bleiben* also gewissermaßen *anonym*. Die Zahl der Teilchen in einem Einteilchenzustand heißt *Besetzungszahl*. Ein Systemzustand wird vollständig durch die Gesamtheit $\{n_1, n_2, \ldots n_k, \ldots\}$ aller Besetzungszahlen fixiert. Die Besetzungszahlen für Bosonen können alle Werte $n = 0, 1, 2, \ldots$ annehmen, die für Fermionen aufgrund des Pauliprinzips nur die Werte $n = 0$ (unbesetzt) und $n = 1$ (besetzt).

Betrachtet man die Besetzungszahlen als die primären meßbaren Größen eines Mehrteilchensystems, so sind gemäß den allgemeinen Prinzipien der Quantentheorie (↑ 2.5.1.4.) die Quantenzustände durch Wahrscheinlichkeitsamplituden $\Phi(n_1\, n_2, \ldots)$ zu charakterisieren. Dabei treten die diskreten Besetzungszahlen als Variable an die Stelle der Ortsvektoren von Teilchen. $|\Phi|^2$ ist die Wahrscheinlichkeit, n_1 Teilchen im Einteilchenzustand 1, n_2 Teilchen im Zustand 2 ... zu finden. Fixierten Besetzungszahlen, wie sie bei wechselwirkungsfreien Teilchen auftreten, entsprechen also spezielle Amplitudenfunktionen Φ, die nur für $n_1 = n_1^0$, $n_2 = n_2^0, \ldots$ betragsmäßig gleich 1 und sonst Null sind; n_k^0 sind dabei die vorgegebenen Besetzungszahlen. Durch eine nicht in dieser Form spezialisierte Funktion Φ wird ein allgemeiner Zustand eines Systems miteinander wechselwirkender Teilchen beschrieben, bei dem die Besetzungszahlen aufgrund von „Stößen" zwischen den Teilchen schwanken. Der auf dieser Grundlage entwickelte *Besetzungszahlformalismus* wird allerdings bei den einfacheren Anwendungen der Quantenmechanik im allgemeinen nicht benutzt.

2.5.5.4. Symmetrisierte Wellenfunktionen. Meist beschreibt man aus identischen Teilchen aufgebaute Systeme trotz der prinzipiellen Schwäche (↑ 2.5.5.3.) dieses Verfahrens durch eine Wellenfunktion $\Phi(\boldsymbol{r}_1, m_1^s; \boldsymbol{r}_2, m_2^s; \ldots)$. Dabei sind die Ortsvariablen $\boldsymbol{r}_n$ durch die Spineinstellungen der Teilchen ergänzt. Diese sind also genau wie die $\boldsymbol{r}_n$ als Variable anzusehen: Ebensowenig wie das Teilchen sich an einem definierten Ort aufhalten muß, braucht es im allgemeinen eine feste Spineinstellung zu besitzen. Im Unterschied zu den kontinuierlich variablen Ortsvektoren durchlaufen die Spinvariablen aber nur diskrete Werte, etwa $m_s = \pm 1/2$ bei Elektronen. Bei einem Teilchen führt dann $\varphi(\boldsymbol{r}, m_s)$ auf die beiden gewöhnlichen Wellenfunktionen $\varphi(\boldsymbol{r}, +1/2) \equiv \varphi_+(\boldsymbol{r})$ und $\varphi(\boldsymbol{r}, -1/2) \equiv \varphi_-(\boldsymbol{r})$, entsprechend bei zwei Teilchen auf vier Funktionen $\Phi_{++}(\boldsymbol{r}_1, \boldsymbol{r}_2)$, Φ_{+-}, Φ_{-+} und Φ_{--}.

Mit dem bloßen Hinschreiben einer Wellenfunktion Φ vom oben erklärten Typ wird eine *formale Numerierung* der Teilchen erzwungen. Daher darf sich wenigstens an den beobachtbaren Größen, an den Wahrscheinlichkeiten und den aus ihnen abgeleiteten Mittelwerten, nichts ändern, wenn die Numerierung willkürlich verändert wird. Mathematisch ist das einer gemeinsamen *Vertauschung von Ortsvektoren und Spinvariablen im Argument von* Φ äquivalent. Bei dieser Operation darf sich die Wellenfunktion höchstens mit einem konstanten komplexen Phasenfaktor $\exp(i\alpha)$ multiplizieren, damit sich die Wahrscheinlichkeiten nicht verändern. Da eine zweimalige Vertauschung der gleichen Variablen die Ausgangsfunktion reproduziert, kann dieser Faktor nur $+1$ oder -1 sein.

Wegen der Identität der Teilchen wären unterschiedliche Symmetriefaktoren ± 1 bei Vertauschung verschiedener Teilchenindizes physikalisch sinnlos. Daraus ergibt sich, daß die Wellenfunktion entweder *vollständig symmetrisch oder vollständig antisymmetrisch* gegenüber der gleichzeitigen Vertauschung der Orts- und Spinvariablen zweier beliebiger Teilchen sein muß. Systeme aus Fermionen werden durch antisymmetrische, Systeme aus Bosonen durch symmetrische Wellenfunktionen beschrieben. Das ist die exakte quantenmechanische Fassung der Gesetzmäßigkeiten bei identischen Teilchen. Das Pauliprinzip in seiner ursprünglichen Formulierung ist darin als Spezialfall enthalten (↑ 2.5.5.5.).

2.5.5.5. Aufbau antisymmetrischer Wellenfunktionen. Für den besonders wichtigen Fall der Fermionen soll die Konstruktion von antisymmetrischen Wellenfunktionen im Rahmen des Modells unabhängiger Teilchen näher untersucht werden. Dazu werden zunächst nur zwei Teilchen mit zwei Spineinstellungen und einem Satz von räumlichen Einteilchenzuständen betrachtet, die unabhängig von der Spineinstellung sein sollen. Die Einteilchenzustände sind $\varphi_k^{(+)}(\boldsymbol{r})$ und $\varphi_k^{(-)}(\boldsymbol{r})$, wobei der Spinindex oben an die Wellenfunktion gesetzt wurde. Es gibt folgende Möglichkeiten, eine antisymmetrische Gesamtfunktion zu konstruieren:

$$k = l: \varphi_k^{(+)}(\boldsymbol{r}_1)\,\varphi_k^{(-)}(\boldsymbol{r}_2) - \varphi_k^{(-)}(\boldsymbol{r}_1)\,\varphi_k^{(+)}(\boldsymbol{r}_2) \quad \text{aus } \varphi_k^{(+)} \text{ und } \varphi_k^{(-)}$$

$$k \neq l \begin{cases} \varphi_k^{(+)}(\boldsymbol{r}_1)\,\varphi_l^{(-)}(\boldsymbol{r}_2) - \varphi_l^{(-)}(\boldsymbol{r}_1)\,\varphi_k^{(+)}(\boldsymbol{r}_2) & \text{aus } \varphi_k^{(+)} \text{ und } \varphi_l^{(-)} \\ \varphi_k^{(-)}(\boldsymbol{r}_1)\,\varphi_l^{(+)}(\boldsymbol{r}_2) - \varphi_l^{(+)}(\boldsymbol{r}_1)\,\varphi_k^{(-)}(\boldsymbol{r}_2) & \text{aus } \varphi_k^{(-)} \text{ und } \varphi_l^{(+)} \\ \varphi_k^{(+)}(\boldsymbol{r}_1)\,\varphi_l^{(+)}(\boldsymbol{r}_2) - \varphi_l^{(+)}(\boldsymbol{r}_1)\,\varphi_k^{(+)}(\boldsymbol{r}_2) & \text{aus } \varphi_k^{(+)} \text{ und } \varphi_l^{(+)} \\ \varphi_k^{(-)}(\boldsymbol{r}_1)\,\varphi_l^{(-)}(\boldsymbol{r}_2) - \varphi_l^{(-)}(\boldsymbol{r}_1)\,\varphi_k^{(-)}(\boldsymbol{r}_2) & \text{aus } \varphi_k^{(-)} \text{ und } \varphi_l^{(-)}. \end{cases}$$

Offenbar ist es nicht möglich, aus zwei gleichen Funktionen, etwa $\varphi_k^{(+)}(\boldsymbol{r}_1)$ und $\varphi_k^{(+)}(\boldsymbol{r}_2)$ eine antisymmetrische Funktion zu bilden. Die beiden Einteilchenzustände müssen sich entweder durch die Spineinstellung oder durch die räumliche Wellenfunktion φ_k oder durch beides unterscheiden. Das ist das Pauliprinzip in seiner einfacheren Formulierung.

Faßt man k, l und die Spineinstellungen zu verallgemeinerten Zustandsindizes α, β zusammen, so lassen sich die eben betrachteten fünf verschiedenen Fälle einheitlich in der Form

$$\Phi \sim \varphi_\alpha(\boldsymbol{r}_1)\,\varphi_\beta(\boldsymbol{r}_2) - \varphi_\beta(\boldsymbol{r}_1)\,\varphi_\alpha(\boldsymbol{r}_2) = \begin{vmatrix} \varphi_\alpha(\boldsymbol{r}_1) & \varphi_\alpha(\boldsymbol{r}_2) \\ \varphi_\beta(\boldsymbol{r}_1) & \varphi_\beta(\boldsymbol{r}_2) \end{vmatrix}$$

darstellen. Auch bei mehr als zwei Teilchen kann Φ als sogenannte *Slaterdeterminante* konstruiert werden. Dabei stehen in verschiedenen Zeilen unterschiedliche Einteilchenzustände, während von Spalte zu Spalte die Teilchenindizes wechseln. Eine Vertauschung zweier Teilchenindizes entspricht der Vertauschung zweier Spalten der Determinante und liefert daher den verlangten Vorzeichenwechsel. Stimmen zwei Einteilchenzustände vollständig, auch in der Spineinstellung, überein, so verschwindet die Determinante wegen der Gleichheit zweier Zeilen.

Bei der Bildung völlig symmetrischer Wellenfunktionen für Bosonen entstehen keinerlei Einschränkungen hinsichtlich der Gleichheit oder Verschiedenheit der Einteilchenzustände.

2.5.5.6. Ausweichen von Fermionen mit gleicher Spineinstellung. Die Antisymmetrie der Wellenfunktion hat Konsequenzen, die über das ursprüngliche Pauliprinzip hinausgehen. Betrachtet man beispielsweise die Wellenfunktion für zwei Teilchen mit gleicher Spineinstellung, so erkennt man, daß Φ für $\boldsymbol{r}_1 = \boldsymbol{r}_2$ verschwindet. Falls $\boldsymbol{r}_1$ und $\boldsymbol{r}_2$ nahe beieinander liegen, wird Φ und damit auch $|\Phi|^2$ und daher die Wahrscheinlichkeit einer solchen Lage der beiden Teilchen zueinander sehr klein sein. Das bedeutet, daß die beiden Teilchen *sich gegenseitig ausweichen*. In der unscharfen Form „wo ein Teilchen ist, kann kein anderes sein" macht das Pauliprinzip diesen Effekt verständlich. Er kann nicht als Folge einer echten Wechselwirkung zwischen den Teilchen verstanden werden, hat aber sehr ähnliche Folgen. So treten im thermodynamischen Verhalten eines idealen Fermigases Effekte auf, wie sie bei einem klassischen realen Gas durch abstoßende Wechselwirkungskräfte zwischen den Gasteilchen bewirkt werden.

Bei Elektronen führt das Ausweichen von Teilchen gleicher Spineinstellung zu einer Verminderung der positiven potentiellen Energie der Coulombschen Abstoßung, da diese Teilchen sich

nicht mehr so nahe kommen. Die Antisymmetrisierung der Wellenfunktion bewirkt damit eine Erniedrigung der Energie. Dieser Energiegewinn wird *Austauschenergie* genannt.
Bei Bosonen hat die Symmetrisierung der Gesamtwellenfunktion qualitativ den gleichen Effekt wie eine Anziehungskraft zwischen den Teilchen eines klassischen Systems. Dies führt dazu, daß sich die Bosonen bevorzugt benachbart aufhalten, daß *lockere Teilchengruppen* gebildet werden. Diese Tendenz geht so weit, daß in einem idealen Bosegas bei tiefen Temperaturen ein Phasenübergang zu einer kondensierten Phase, dem sogenannten *Bosekondensat*, stattfindet. Diese Möglichkeit wurde zuerst von EINSTEIN gesehen.

2.5.6. Zeitabhängigkeit

2.5.6.1. Zeitabhängige Schrödingergleichung. Die Wellenfunktion eines Zustands mit genau bestimmter Energie besitzt entsprechend der Planckschen Beziehung die Zeitabhängigkeit $\varphi_E(t) = \varphi_E \exp(-iE_n t/\hbar)$, die charakteristisch für die stationären Zustände ist (↑ 2.5.4.1.). Im allgemeinen können viele Energiewerte einschließlich der des kontinuierlichen Spektrums mit den ihnen zugeordneten Wahrscheinlichkeiten auftreten. Die Wellenfunktion $\psi(t)$ eines solchen Zustands stellt eine Linearkombination von Funktionen $\varphi_E(t)$ dar. Da die Energieeigenfunktionen φ_E der Energieeigenwertgleichung $\hat{H}\varphi_E = E\varphi_E$ genügen, gilt $-(\hbar/i)\,\partial\varphi_E(t)/\partial t = \varphi_E E \mathrm{e}^{-iEt/\hbar} = \hat{H}\varphi_E \mathrm{e}^{-iEt/\hbar} = \hat{H}\varphi_E(t)$ (↑ 2.5.2.1.). In dieser Form tritt der jeweilige Energiewert E nicht mehr in der Gleichung selbst auf. Die *zeitabhängige Schrödingergleichung*

$$(5)\qquad -\frac{\hbar}{i}\frac{\partial\psi}{\partial t} = \hat{H}\psi$$

bestimmt daher den zeitlichen Ablauf einer beliebigen Linearkombination von Funktionen $\varphi_E(t)$ und damit einer beliebigen Wellenfunktion ψ überhaupt.
Da die Schrödingergleichung hinsichtlich der Zeitabhängigkeit von erster Ordnung ist, ist der gesamte Zeitablauf $\psi(t)$ eindeutig festgelegt, wenn die Wellenfunktion zu irgendeiner Anfangszeit bekannt ist.
Die Schrödingergleichung für *ein* Teilchen, für den Hamiltonoperator $\hat{H} = (-\hbar^2/2m)\Delta + V(\boldsymbol{r})$, stimmt mit der entsprechenden Gleichung für Materiewellen (↑ 1.2.3.) überein. In diesem Zusammenhang wurde sie von SCHRÖDINGER ursprünglich aufgestellt. Sie bleibt aber auch für beliebig komplizierte Teilchensysteme gültig, solange diese nichtrelativistisch behandelt werden können. Die Wellenfunktion $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, t)$ hängt dann außer von der Zeit von den Ortsvektoren aller Teilchen ab, im allgemeinen außerdem von hier nicht angegebenen Spinvariablen (↑ 2.5.5.4.).

2.5.6.2. Zeitabhängigkeit von Mittelwerten, Erhaltungsgrößen. Die wie üblich als mittlere Meßwerte definierten quantenmechanischen Mittelwerte $\bar{A}$ können nach einer integralen Vorschrift [↑ 2.5.(3)] aus der Wellenfunktion ψ berechnet werden. Mittels der Schrödingergleichung kann eine einfache Aussage über die erste *Zeitableitung der Mittelwerte* abgeleitet werden. Sie lautet

$$(6)\qquad \frac{\mathrm{d}\bar{A}}{\mathrm{d}t} \equiv \dot{\bar{A}} = \frac{i}{\hbar}\,\overline{(\hat{H}\hat{A} - \hat{A}\hat{H})}.$$

Dabei definiert die antisymmetrische Bildung in der Klammer, die sogenannte *Vertauschung der Operatoren* $\hat{H}$ und $\hat{A}$, einen neuen Operator. Die durch den Querstrich symbolisierte Mittelwertbildung ist genau so wie bei $\bar{A}$ auszuführen.
Von besonderem Interesse sind diejenigen physikalischen Größen, deren Operatoren $\hat{A}$ mit dem Energieoperator vertauschbar sind, d. h., daß es auf die Reihenfolge $\hat{H}\hat{A}$ oder $\hat{A}\hat{H}$ bei der Anwendung beider Operatoren auf eine beliebige Wellenfunktion ψ nicht ankommt. Die quantenmechanischen Mittelwerte solcher Größen sind, als Folge der angegebenen allgemeinen Relation für $\dot{\bar{A}}$, zeitlich konstant, und zwar für jede mögliche Wellenfunktion, gewissermaßen für jeden möglichen quantenmechanischen „Bewegungsablauf". Darüber hinaus läßt sich zeigen, daß nicht nur der Mittelwert, sondern die gesamte Wahrscheinlichkeitsverteilung für die einzelnen Meßwerte konstant bleiben. Das ist genau das, wodurch man eine *Erhaltungsgröße in der Quantentheorie* charakterisieren muß.
Die Vertauschbarkeit von $\hat{A}$ mit $\hat{H}$ ist in zwei einfachen, aber wichtigen Fällen stets gewährleistet, nämlich für $\hat{A} = 1$ und $\hat{A} = \hat{H}$, da jeder Energieoperator mit dem trivialen Operator 1 und mit sich selbst vertauschbar ist. Zu $\hat{A} = 1$ gehört als Mittelwert das *Normierungsintegral* $\int \mathrm{d}V\psi^*\psi$ der Wellenfunktion. Eine Wellenfunktion, die zu irgendeiner Zeit normiert war, behält diese Eigenschaft im Zeitablauf bei. Dieser Umstand ist entscheidend für die Wahrscheinlichkeitsinterpretation der Theorie, denn die Normierung von ψ folgte direkt aus der Auffassung von $|\psi|^2$ als Wahrscheinlichkeitsdichte für Ortsmessungen (↑ 2.5.1.5.).
Zu $\hat{A} = \hat{H}$ gehört als Mittelwert $\bar{H}$ die *mittlere Energie* $\bar{E}$ des Systems. Diese bleibt, wie überhaupt die gesamte Energieverteilung, zeitlich ungeändert.
Größen, die bei der klassischen Bewegung Erhaltungsgrößen sind, bleiben es auch in der Quantenmechanik, d. h., ihre Operatoren sind mit $\hat{H}$ vertauschbar. Überhaupt gelten für die Zeitableitungen der Mittelwerte ähnliche Beziehungen wie für die klassischen Größen. Betrachtet man beispielsweise den mittleren Ort $\bar{\boldsymbol{r}}$ und den mitt-

leren Impuls $\bar{\boldsymbol{p}}$ für ein *Wellenpaket* $\psi(r, t)$, so führt die Auswertung der entsprechenden Operatorvertauschungen auf die *Ehrenfestschen Sätze* $\dot{\bar{\boldsymbol{r}}} = \frac{\bar{\boldsymbol{p}}}{m}$ und $m\ddot{\bar{\boldsymbol{r}}} = \dot{\bar{\boldsymbol{p}}} = -\overline{\text{grad } V} = \overline{\boldsymbol{F}(\boldsymbol{r})}$. Dabei ist $\boldsymbol{F}(\boldsymbol{r})$ die ortsabhängige Kraft. Die klassische Newtonsche Bewegungsgleichung gilt gewissermaßen weiterhin „im Mittel".

2.5.6.3. Übergangswahrscheinlichkeiten. Äußere Störungen können die Energieverteilung ändern. Insbesondere bleibt ein System dann nicht mehr in einem stationären Zustand des ungestörten Systems. Befindet sich das System beispielsweise im Grundzustand und wirkt dann kurzzeitig eine *Störung*, etwa durch ein vorbeifliegendes Teilchen, so besteht hinterher eine endliche Wahrscheinlichkeit dafür, das System in einem angeregten Zustand anzutreffen. Diese sogenannten Übergangswahrscheinlichkeiten haben für die verschiedenen Anregungszustände verschiedene Werte. Möglich sind auch Störungen, die zu einer Zeit „eingeschaltet" werden und dann bestehen bleiben; eine plötzliche Änderung des Atomkerns durch einen spontanen Umwandlungsprozeß wirkt auf diese Weise auf die Elektronen des Atoms. Die Übergangswahrscheinlichkeiten beziehen sich dann auf die Energiezustände des „neuen" Atoms mit dem „neuen" Kern.

Schließlich gibt es Störungen wie die Wechselwirkung der Elektronen mit dem elektromagnetischen Strahlungsfeld, die dauernd wirken und so beschaffen sind, daß Zustände, die bei ihrer Vernachlässigung exakt stationär wären, tatsächlich nur eine endliche Lebensdauer haben. Solche Zustände heißen *quasistationäre Zustände*. Sie können offenbar nur als Zwischenzustände von Streu- bzw. Reaktionsprozessen auftreten. Sie führen dort zu *Resonanzen* in den Wirkungsquerschnitten (↑ 2.4.3.5.). Zwischen der Energiebreite ΔE dieser Resonanzen und der Lebensdauer τ der zerfallenden Zustände besteht die allgemeine Beziehung $\tau \, \Delta E \approx \hbar$. Falls die Zustände genügend langlebig sind, können sie auch unabhängig von ihren Bildungsprozessen beobachtet werden; das trifft z. B. für radioaktive Atomkerne zu.

3. Elementarteilchen und Atomkerne

3.1. Teilchenbeschleuniger

Elementarteilchen bzw. Atomkerne können nur miteinander wechselwirken, wenn sie eine hinreichend große kinetische Energie besitzen. Bei der Wechselwirkung zwischen Atomkernen ist die Coulombabstoßung zu überwinden; bei Reaktionen zwischen Elementarteilchen entstehen neue Teilchen, z. B. Mesonen, Pionen, Baryonen, sofern die entsprechenden Energieschwellen überschritten werden.
Beschleunigt werden geladene Teilchen mit Hilfe eines elektrischen Felds (↑ 2.1.6.5.), welches als statisches Feld, elektrisches Wechselfeld oder als von einem veränderlichen Magnetfeld erzeugtes Wirbelfeld vorliegen kann. Die erhaltene kinetische Energie E_k ist dabei nur durch die Teilchenladung q und die durchlaufene Potentialdifferenz U bestimmt: $E_k = qU$. Ihre Einheit ist das Elektronenvolt (eV).
Die verschiedenen Massen der Teilchen sowie die unterschiedlichen Endenergien gestatten kein einheitliches Beschleunigerprinzip. Das Elektron mit seiner geringen Ruhmasse ($mc^2 = 511$ keV) erreicht bereits unterhalb 1 MeV relativistische Geschwindigkeiten ($v \approx c$), bei schweren Teilchen geschieht das erst im GeV-Bereich. In der sechzigjährigen Geschichte der Beschleunigertechnik haben sich deshalb einige prinzipielle Entwicklungslinien herausgebildet.

3.1.1. Linearbeschleuniger

3.1.1.1. Elektrostatische Beschleuniger. Die in einer Ionenquelle erzeugten Ionen treten in die evakuierte Beschleunigungsstrecke zwischen zwei Hochspannungselektroden. Weitere Elektroden sorgen für eine gleichmäßige Verteilung des Potentials über die Beschleunigungsstrecke sowie für die ionenoptische Führung der Teilchen. Die notwendige Hochspannung wird verschiedenartig erzeugt.

3.1.1.1.1. *Kaskadengenerator.* Er basiert auf dem von GREINACHER 1921 angegebenen Prinzip der Spannungsvervielfachung. Dabei wird eine Wechselspannung von etwa 100 kV in eine Gleichspannung von mehreren MV umgewandelt. Mit einem derartigen Generator wurde 1932 durch COCKROFT und WALTON die erste Kernreaktion mit künstlich beschleunigten Teilchen $^7Li + {}^1H \rightarrow 2\,{}^4He$ hervorgerufen. Kaskadengeneratoren dienen heute als Vorstufen für Beschleuniger höchster Energien sowie als Neutronengeneratoren (↑ 3.4.4.2.).

3.1.1.1.2. *Van-de-Graaff-Generator.* In ihm erfolgt der Ladungstransport zur Hochspannungselektrode durch ein Gummiband oder eine Kette aus isolierten Gliedern. Hohe Spannungen oberhalb 1 MV erfordern aus Isolationsgründen die Installation der Anlage in einem Druckkessel. Heute werden Feldstärken bis zu 30 MVm^{-1} beherrscht.

3.1.1.1.3. *Tandem-Van-de-Graaff-Generator.* Durch Umladung der Ionen läßt sich in diesem Generator die Spannung U für den Beschleunigungsprozeß zwei- bzw. mehrfach ausnutzen (Abb. 3.1.1-1). Dabei erfolgt zunächst die Beschleunigung von negativ geladenen Ionen vom Erdpotential aus zu einer positiv geladenen Hochspannungselektrode. Die Umladung in positiv geladene Ionen findet in einer speziellen Folie oder einer Gaskammer im Innern der Elektrode durch Abstreifen eines Teils der Hüllenelektronen (Strippertarget) statt, so daß eine weitere Beschleunigung zum Erdpotential hin erfolgt.

3.1.1.2. HF-Linearbeschleuniger. Höhere Energien lassen sich durch Vervielfachung der Beschleunigungsstrecken erzielen, an denen eine hochfrequente Beschleunigungsspannung angelegt ist. Die Länge der Strecken und die jeweilige Geschwindigkeit der Teilchen sind aufeinander abgestimmt. Bei geeignet gewählter Frequenz und HF-Spannung wirkt die elektrische Feldstärke nur in Beschleunigungsrichtung (Phasenbedingung). Die Methode wurde 1928

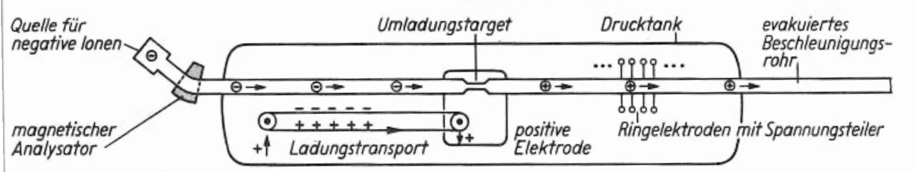


Abb. 3.1.1-1 Prinzipielle Wirkungsweise des Tandem-Van-de-Graaff-Generators

von WIDERÖE vorgeschlagen. Wirkungsvolle Beschleuniger wurden jedoch erst mit der Verfügbarkeit der UKW-Technik und der Beherrschung der Physik der Hohlraumresonatoren gebaut.

3.1.1.2.1. *Beschleuniger für Elektronen und Ionen mit stehenden Wellen.* Die Teilchen werden hier zwischen Driftröhren beschleunigt. Eine zeitliche Stabilisierung der Phasenlage über viele Beschleunigungsschritte hinweg erfolgt automatisch (Prinzip der Phasenfokussierung). Räumlich werden die Teilchen in den Driftstrecken durch ionenoptische Einrichtungen (z. B. magnetische Linsen) zusammengehalten. Die technische Grenze für alle Linearbeschleuniger ist bei etwa 1,5 $\mathrm{MeVm^{-1}}$ erreicht. Ihre häufigste Anwendung ist die Vorbeschleunigung von Ionen auf etwa 100 MeV für die Beschleuniger höchster Energien (↑ 3.1.2.4.).

3.1.1.2.2. *Beschleuniger mit fortschreitenden Wellen.* Die Beschleunigung der Teilchen erfolgt an der Flanke einer Wanderwelle, man spricht vom Wellenreiterprinzip. Durch geometrische Gestaltung des Wellenleiters wird die Phasengeschwindigkeit der Wanderwelle an die Teilchengeschwindigkeit angepaßt. Voraussetzung ist eine Teilchengeschwindigkeit bereits nahe der Lichtgeschwindigkeit. Deshalb ist das Prinzip besonders für Elektronen geeignet. Beispiele sind der 2-GeV-Beschleuniger in Charkow und der 20-GeV-Beschleuniger in Stanford (SLAC) mit einer Länge von 3 km. Bei letzterem ist eine Rekonstruktion mit supraleitendem Wellenleiter auf eine Endenergie von 100 GeV vorgesehen.

3.1.1.3. Kollektivbeschleuniger. Eine andere Variante des Wellenreiterprinzips stellt die Mitführung von positiv geladenen Ionen durch eine beschleunigte Elektronenwolke dar. Bei Beschleunigung der Elektronen in den relativistischen Bereich werden die Ionen auf die gleiche Geschwindigkeit mitgezogen und erlangen damit gegenüber den Elektronen eine wesentlich höhere kinetische Energie. So könnten mit der Beschleunigung von Elektronen

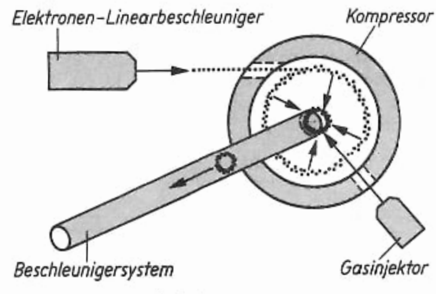


Abb. 3.1.1-2 Prinzipielle Wirkungsweise des Kollektivbeschleunigers

auf 20 MeV auf einer Länge von 15 m Protonen auf eine Energie von 1 GeV gebracht werden. Das Prinzip wurde 1956 von BUDKER und WEKSLER angegeben. Seitdem arbeiten Laboratorien in Dubna, München und Maryland (USA) an der Realisierung eines derartigen Beschleunigers. Bisher ist es gelungen, stabile Elektronenpakete in Form eines Rings von einigen cm Durchmesser zu erzeugen, indem relativistisch beschleunigte Elektronenstrahlen in einem Magnetfeld aufgespult und bei steigendem Feld komprimiert wurden. Hauptanliegen der gegenwärtigen Entwicklungsschritte ist die Beladung des Elektronenrings mit Ionen und seine Beschleunigung. Ein nach diesem Prinzip arbeitender Kollektivbeschleuniger würde Teilchenimpulse extrem kurzer Dauer, d. h. im ps-Bereich liefern (Abb. 3.1.1-2).

3.1.2. Kreisbeschleuniger

Die Ablenkung geladener bewegter Teilchen mit dem Impuls p in einem Magnetfeld B (↑ 2.1.6.6.) gestattet es, die Dimensionen des Beschleunigers auch bei steigenden Endenergien gegenüber linearen Beschleunigern zu reduzieren. Allgemein gilt streng relativistisch für das Produkt aus Feldstärke B und Krümmungsradius ϱ die Relation $B\varrho = p/q$, wobei q die elektrische Ladung des Teilchens ist.

3.1.2.1. Klassisches Zyklotron. Ein nichtrelativistisches Teilchen der Masse m bewegt sich im homogenen Magnetfeld B auf einer Kreisbahn mit der von seiner Energie bzw. Geschwindigkeit unabhängigen Kreisfrequenz $\omega = qB/m$. Das gestattet ein Beschleunigungsprinzip mit fester Frequenz (Abb. 3.1.2-1). Die Beschleunigung von Teilchenpaketen erfolgt phasenrichtig durch eine HF-Spannung, welche an zwei D-förmigen Elektroden in einer Vakuumkammer zwischen den Polschuhen eines Magneten anliegt. Diese Duanten („Dees") sind Teile eines Resonators, welcher durch einen leistungsfähigen HF-Sender angeregt wird. Die Amplitude der Beschleunigungsspannung beträgt etwa 100 kV.

Die Grenze des klassischen Festfrequenzzyklotrons liegt für Protonen bei etwa 10 MeV. Eine stabile Führung des Strahls über viele Umläufe ist nur bei einem nach außen abfallenden Feld möglich. Nur in diesem Fall führen die beschleunigten Teilchen sogenannte Betatronschwingungen um die Soll-Strahllage aus. Damit verbunden ist aber nach einer Reihe von Umläufen die Verletzung der Frequenzbedingung, das Teilchenpaket wird nicht weiter beschleunigt und verbleibt auf dem sogenannten Synchronkreis.

3.1.2.2. Mikrotron. Elektronen können im Zyklotron nicht beschleunigt werden, da sehr

bald die durch den relativistischen Massenzuwachs bedingte Grenze erreicht ist. Ein phasenrichtiger Durchgang der Elektronen in einem homogenen Magnetfeld ist jedoch möglich, wenn die Teilchen in jedem Beschleunigungsschritt den Energiezuwachs 0,51 MeV (Ruhenergie des Elektrons) erhalten. In

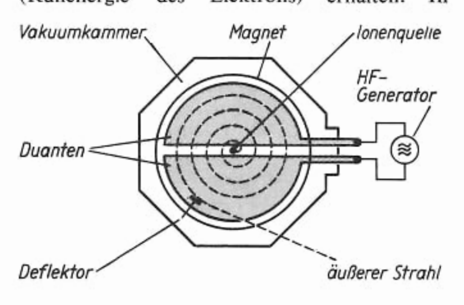


Abb. 3.1.2-1 Prinzipielle Wirkungsweise des Zyklotrons

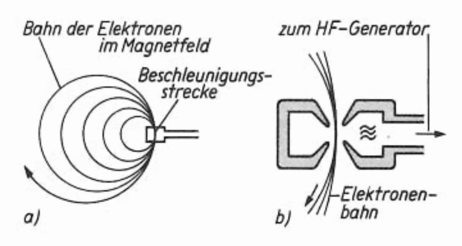


Abb. 3.1.2-2 Prinzipielle Wirkungsweise des Mikrotrons

diesem Fall beträgt die Umlaufzeit immer ein Vielfaches der Grundperiode $\tau = 2\pi m/eB$. Bei einer Beschleunigungsspannung mit der Kreisfrequenz $\omega = 2\pi/\tau$ erfolgt in jedem Schritt eine phasenrichtige Beschleunigung. Die Grenzen des Beschleunigers liegen bei etwa 10 MeV. Sie sind bedingt durch die bereits merklich einsetzende Strahlungsdämpfung (Synchrotronstrahlung proportional E^4/ϱ^2). Das Prinzip wird zur Erzeugung harter Röntgen- bzw. γ-Strahlung benutzt (Abb. 3.1.2-2).

3.1.2.3. Isochronzyklotron. Für die phasengetreue Beschleunigung von Ionen bei fester Frequenz muß das Magnetfeld des Zyklotrons mit der Energie, also mit wachsendem Radius, entsprechend $B = \gamma B_0$ ansteigen (B_0 Magnetfeld im Zentrum, $\gamma = 1/\sqrt{1-\beta^2}$ relativistischer Faktor des Massenzuwachses mit $\beta = v/c$). Das widerspricht der Forderung nach einem schwachen Feldabfall nach außen zum Zweck der Stabilisierung der Teilchenbahnen. Da obige Bedingung jedoch für ein längs der Teilchenbahn gemitteltes Magnetfeld gilt, kann eine Stabilisierung des Strahls durch ionenoptische Prinzipien herbeigeführt werden.

Bei der *Thomasfokussierung* wird dazu eine räumliche Variation des Magnetfelds benutzt. Zu diesem Zweck haben die Polschuhe des Zyklotrons eine sektorförmige Struktur, so daß sich während eines Umlaufs mehrere stärkere und schwächere Feldbereiche abwechseln (azimutalvariierendes Feld). Dieses Prinzip wird bei geringen Annäherungen an relativistische Geschwindigkeiten ($\beta \approx 0{,}2$) benutzt. Eine stärkere Stabilisierung des Strahls wird erreicht, wenn die Feldsektoren zusätzlich eine spiralförmige Geometrie haben (Abb. 3.1.2-3) oder durch eine vollständige Trennung der Feldsektoren (Sektorfokussierung).

Isochronzyklotrone sind durch gute Strahleigenschaften charakterisiert. Die Grenzen der Beschleunigung sind durch die Größe der Anlage und das Gewicht des Eisenjochs bei etwa 1 GeV für Protonen gegeben. Die Mehrzahl der auf der Welt existierenden etwa 75 Isochronzyklotrone erzeugt Protonen im Energiegebiet zwischen 20 und 100 MeV (Kompaktzyklotrone), bzw. schwere Ionen mit entsprechend höheren Energien $E = E_{\text{Proton}}q^2/A$ (q Ladungszahl, A Massenzahl der Ionen). Das z. Z. größte Isochronzyklotron mit einem Polschuhdurchmesser von 4 m befindet sich im VIK Dubna (Tafel 3). Es beschleunigt schwere Ionen auf die Endenergie $E = 850q^2/A$ MeV. Noch größere Anlagen sind in Sektorbauweise ausgeführt. Hierin können relativ große Teilchenströme beschleunigt werden, mit denen durch Sekundärreaktionen Mesonen in genügender Menge erzeugt werden können (Mesonenfabriken), welche ihrerseits für eine Reihe von Anwendungsfällen benötigt werden. Hohe Endenergien lassen sich bei wesentlich kleineren Abmessungen mit supraleitenden Magneten erzeugen. Entsprechende Projekte sind an der Michigan State University (USA) und in Chalk River (Kanada) in Vorbereitung.

3.1.2.4. Synchrozyklotron. Eine phasengetreue Beschleunigung ist auch durch Verringerung der Frequenz im konventionellen Zyklotronfeld während des Beschleunigungsprozesses mög-

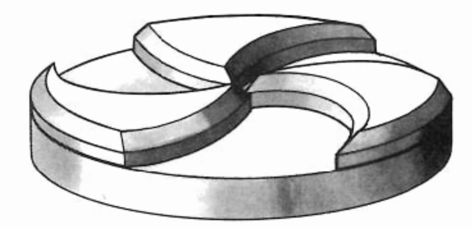


Abb. 3.1.2-3 Profil der Oberfläche eines Polschuhs im Isochronzyklotron

lich. Die von WEKSLER und MCMILLAN 1945 gefundene Phasenstabilität der Teilchen auf dem Synchronkreis des Festfrequenzzyklotrons gestattet, den Synchronkreis durch Frequenzverringerung im Laufe des Beschleunigungsprozesses langsam nach außen zu verschieben und somit die Energie über die dem Festfrequenzzyklotron eigenen Grenzen zu steigern. Bei relativ geringer Beschleunigungsspannung laufen die Teilchen bis zu 10000mal um. Die Hochfrequenzspannung wird durch einen rotierenden Kondensator mit einer Frequenz von etwa 100 Hz moduliert. Beschleunigt wird in der Modulationsperiode nur ein Teilchenpaket. Frequenzmodulierten (FM-)Zyklotronen ist durch ihre Größe und damit das notwendige Eisen- und Kupfergewicht eine Energiegrenze bei etwa 1 GeV gesetzt. Das Synchrozyklotron des VIK Dubna besitzt einen Polschuhdurchmesser von 6 m bei einem Gewicht des Eisenjochs von 7000 t. Es beschleunigt Protonen auf eine Endenergie von 680 MeV. Zyklotrone dieses Typs stellen die erste Generation von Mesonenfabriken dar. Sie sind heute durch das moderne Prinzip des Isochronzyklotrons (↑ 3.1.2.3.) weitgehend abgelöst.

3.1.2.5. Synchrotron. Die weitere Erhöhung der Energien wurde durch die Führung der Teilchen in ringförmig angeordneten Magneten, d. h. bei gleichbleibendem Radius der Umlaufbahnen, möglich. Notwendig ist hierbei, sowohl das Magnetfeld als auch die Frequenz während der Beschleunigung eines Teilchenpakets zu verändern.

Tab. 3.1.2-4 Beschleuniger für höchste Energien

1. Generation: Protonensynchrotrons mit schwacher Fokussierung (Strahlquerschnitt etwa 1 m²; q Durchmesser, E Energie)

Inbetriebnahme	Ort	Bezeichnung	q in m	E in GeV
1953	Brookhaven	Cosmotron	18	3
1954	Berkeley	Bevatron	38	6
1957	VIK Dubna	Synchrophasotron	72	10
2. Generation: Protonensynchrotrons mit starker Fokussierung (Strahlquerschnitt etwa 100 cm²)				
1959	CERN	CPS	200	28
1960	Brookhaven	AGS	257	33
1967	Serpuchow	U 70	472	76
1971	Batavia (USA)	Fermilab Accelerator	2000	500
1976	CERN	SPS	2200	400

3.1.2.5.1. *Protonensynchrotron.* Die erste Generation dieser ringförmigen Beschleuniger wurde in ihrer Entwicklung durch die großen Strahlquerschnitte und den damit verbundenen hohen Aufwand bei der Errichtung der Magnete begrenzt. Erst die Einführung von wechselnden Magnetfeldgradienten längs der Bahn führte zu einer stärkeren Fokussierung der Teilchenbahnen und gestattete eine weitere Vergrößerung der Anlagen bei geringeren Strahlquerschnitten (Tab. 3.1.2-4; Tafeln 4/5).

Einen Schritt zur weiteren Erhöhung der effektiven Teilchenenergien stellte die Schaffung von gegenläufigen oder sich kreuzenden Speicherringen dar, welche es ermöglichen, Teilchenströme aufeinanderzuschießen und somit selbst als Target zu verwenden. Der Vorteil besteht darin, daß damit für die Reaktion die Summe der Teilchenenergien wirklich zur Verfügung steht. Die geringen Wirkungsquerschnitte bei derartigen Experimenten werden zum Teil dadurch kompensiert, daß es möglich ist, in den Speicherringen große Teilchenströme bis zu mehreren A anzusammeln. Derartige Experimente sind auf sehr konzentrierte Teilchenbündel, d. h. also auf kleine Strahlquerschnitte, angewiesen. Zu diesem Zweck wurden in Nowosibirsk und bei CERN Methoden zur Verringerung der Impulsunschärfe im Strahl entwickelt.

Vor wenigen Jahren begannen Entwicklungen, durch supraleitende Erregerspulen die maximale Feldstärke der Ablenkmagnete in Ringbeschleunigern zu erhöhen. Außer einer Verdopplung der Energie gegenüber konventionellen Magneten vergleichbarer Dimensionen erfolgt dabei eine beträchtliche Einsparung an Elektroenergie für die Erregung der Magnete.

Beschleunigerkomplexe für Experimente bei höchster Energie bestehen aus mehreren Vorbeschleunigern, dem Hauptbeschleuniger, einem Speicherring und einem Strahlleitungssystem, welches die beschleunigten Strahlen zum Experiment führt. Eines der zur Zeit gewaltigsten Projekte wird in Serpuchow bei Moskau erarbeitet. Es handelt sich um einen Speicherringkomplex mit supraleitendem Synchrotron einer Endenergie von 3 TeV. Der Supraleiterring wird sich in einem unterirdischen Tunnel mit einer Länge von 19 km befinden (Tafeln 4/5).

3.1.2.5.2. *Elektronensynchrotron.* Elektronen werden mit dem gleichen Prinzip, jedoch infolge der Strahlungsdämpfung nur auf Energien bis zu etwa 20 GeV beschleunigt. Ihre Vorbeschleunigung kann im selben Ring durch Induktionsbeschleunigung nach dem Betatronprinzip erfolgen. Da Elektronen sehr schnell relativistische Energien erreichen, bleibt die Frequenz

konstant. Das größte Elektronensynchrotron befindet sich in Hamburg (DESY). In ihm werden Elektronen in einem Ring von 50 m Durchmesser auf 7,5 GeV beschleunigt. Ein Speicherring (PETRA) mit einem Durchmesser von 400 m ermöglicht Experimente mit gegenläufigen Elektronen und Positronen bei Endenergien von 19 GeV. Pläne für die achtziger Jahre sehen Anlagen mit Beschleunigungsenergien von mehr als 100 GeV vor.

3.1.2.6. Betatron. Im Betatron geschieht der Beschleunigungsvorgang nach dem Transformatorprinzip (elektrisches Wirbelfeld als Folge der zeitlichen Änderung des magnetischen

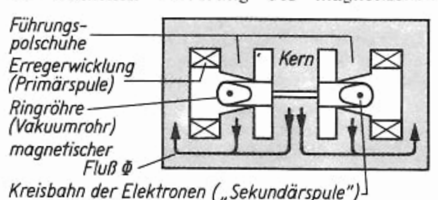


Abb. 3.1.2-5 Prinzipskizze des Betatrons

Felds), wobei die Sekundärwicklung des Transformators aus dem Strom der geladenen Teilchen in einem ringförmigen Vakuumgefäß besteht (Abb. 3.1.2-5). Die Betriebsfrequenz beträgt 50 bis 1 000 Hz. Die Forderung nach der entsprechenden Krümmung der Teilchenbahn führt zum Zusammenhang $B_s = \frac{1}{2}\bar{B}$ zwischen dem magnetischen Beschleunigungsfeld $\bar{B}$ (mittleres Feld innerhalb des beschleunigten Teilchenstroms) und dem Führungsfeld B_s (Wideröebedingung). Die Teilchen führen Betatronschwingungen um den Sollkreis aus, sofern das Feld den ionenoptischen Bedingungen gehorcht (Feldabfall nach außen). Für Ionen eignet sich das Prinzip nicht. Die Grenze der Energie ist durch die Strahlungsdämpfung gegeben und liegt theoretisch bei etwa 500 MeV. Praktisch werden jedoch Betatrone nur bis zu Energien von 50 MeV für die Erzeugung intensiver γ-Strahlung in der Technik und Medizin verwendet (Tafel 11).

3.1.3. Anwendungen der Beschleuniger

Bis 1960 waren die Beschleuniger fast ausschließlich das Instrument der Kern- und der Elementarteilchenphysiker. Danach fanden sie jedoch auch in Bereichen anderer Grundlagenwissenschaften, in der angewandten Forschung und sogar in der Produktionssphäre eine immer breitere Anwendung. Dabei zeichnen sich besonders die drei folgenden Gebiete ab:

die Nutzung der Synchrotronstrahlung aus Elektronenspeicherringen, welche vielfach ausschließlich für diesen Zweck errichtet werden;

die Nutzung von beschleunigten schweren Ionen: Um für die Ionen im Beschleunigungsprozeß einen möglichst hohen Ladungszustand zu erreichen, werden vielfach Beschleunigerkopplungen mit zwischengeschalteten Umladungssektionen (Strippertargets) aufgebaut. Da die Anwendungsmöglichkeiten außerordentlich vielfältig sind, werden in verschiedenen Staaten Applikationszentren mit einem oder mehreren Schwerionenbeschleunigern errichtet;

die Nutzung von Protonen oder Deuteronen für medizinische Zwecke. In einer Reihe von Kliniken gibt es bereits eigene Beschleuniger mit ausgesprochen medizin- und biologieorientierten Programmen der Forschung und Anwendung. Im Vordergrund stehen dabei die Erzeugung von kurzlebigen Radioisotopen für diagnostische Zwecke und die Krebstherapie mittels schneller Neutronen.

3.2. Elementarteilchen

Die Physik der Elementarteilchen ist im wesentlichen eine Physik hoher Energien. Die überwiegende Anzahl von Teilchen tritt unter natürlichen irdischen Bedingungen nicht auf. Um diese Teilchen überhaupt erst im Laboratorium zu erzeugen, müssen das ursprünglich vorliegende bzw. relativ leicht zu gewinnende Elektron und Proton auf Energien im Gigaelektronvoltbereich beschleunigt und auf einen materiellen Target oder einen zweiten Teilchenstrahl geschossen werden. Die Erforschung der zwischen Elementarteilchen ablaufenden Streu- und Umwandlungsprozesse und die Suche nach neuen Teilchen mit größerer Ruhmasse erfordert den Bau immer leistungsfähigerer Teilchenbeschleuniger (↑ 3.1.).

3.2.1. Nachweis und Eigenschaften von Elementarteilchen

3.2.1.1. Entwicklung und Inhalt der Elementarteilchenphysik. Unter dem Begriff Elementarteilchen faßt man allgemein die Gesamtheit der subatomaren Teilchen zusammen, deren bekannteste Vertreter Photon, Elektron, Proton und Neutron sind. Solange nur diese wenigen Teilchen bekannt waren, durch deren Zusammenwirken Struktur und Verhalten der Atome und Atomkerne mit Ausnahme des β-Zerfalls (↑ 3.4.1.) vollständig verstanden werden können, schien die Situation einfach.

Die Vielfalt der Elementarteilchen deutete sich zuerst in der kosmischen Höhenstrahlung an, wo das Positron (1932), das Myon und das Pion (1947) beobachtet wurden. Eine systematische Erforschung dieser Teilchen begann mit dem

Einsatz von Beschleunigern; dabei wurde auch eine Vielzahl bis dahin nicht bekannter Teilchen gefunden. Nachdem diese Entwicklung schon zu einem Abschluß gekommen zu sein schien, stellte die Entdeckung der sogenannten neuen Teilchen eine außerordentliche Überraschung dar, die zu neuen physikalischen Vorstellungen Anlaß gab.
Die Elemantarteilchenphysik unterscheidet sich hinsichtlich der umgesetzten Energien deutlich von anderen Gebieten der Mikrophysik: sind es bei Atomen und Molekülen eV und bei Atomkernen MeV, so handelt es sich hier um GeV, ja neuerdings um Hunderte von GeV. Einem großen Teilchenimpuls p entspricht, im Sinne des Welle-Teilchen-Dualismus (↑ 1.2.4.) nach der de-Broglie-Beziehung, die sehr kleine Wellenlänge $\lambda = h/p$, die wie in der Optik eine hohe Auflösung ermöglicht. Der Einsatz höherer Energien erlaubt demnach das Erfassen immer kleinerer Strukturen. Mit dem Begriff der Elementarteilchen verbindet sich also keineswegs die Vorstellung kleinster, nicht weiter auflösbarer Objekte.
Die meisten Teilchen sind instabil, sie zerfallen und wandeln sich dabei in andere Teilchen um. Auch bei der Wechselwirkung von Teilchen miteinander entstehen im allgemeinen völlig neue Teilchen, darunter bei ausreichend hohen kinetischen Energien der Stoßpartner auch schwerere Teilchen als ursprünglich vorhanden waren. Diese Prozesse können im allgemeinen nicht als bloße Umverteilung von Grundbausteinen aufgefaßt werden wie etwa die Kernreaktionen (↑ 3.4.2.). Die Elementarprozesse der Teilchenphysik sind die der Erzeugung und Vernichtung von Teilchen. Die Gesetze der Teilchenphysik werden durch die Untersuchung von Streu-, Erzeugungs- und Zerfallsprozessen studiert. Die Ergebnisse dieser Experimente werden quantitativ durch die Angabe von Wirkungsquerschnitten (↑ 2.4.3.5.), Lebensdauern und Verzweigungsverhältnissen ausgedrückt. Die Untersuchungen zeigen, daß nicht alle mit den bekannten Erhaltungssätzen für Energie, Impuls und elektrische Ladung verträglichen Prozesse beobachtet werden. Es ist möglich, den Teilchen gewisse diskrete Quantenzahlen als für sie charakteristische Größen so zuzuordnen, daß die möglichen Prozesse durch Erhaltungssätze für diese neuen Größen von den nicht möglichen unterschieden werden. Außerdem zeigt die quantitative Auswertung, daß die Wechselwirkungsprozesse zwischen Elementarteilchen von sehr verschiedener „Intensität" sind, charakterisiert durch die Größe der Wirkungsquerschnitte bzw. die Kleinheit der Lebensdauern instabiler Teilchen. Diese Erfahrungen haben zur Klassifikation der Teilchen in Hadronen und Leptonen geführt und zur Unterteilung der Wechselwirkungen in starke, elektromagnetische und schwache Wechselwirkungen.
Die Elementarteilchen unterliegen nach heutiger Kenntnis voll und ganz den Gesetzen von Quanten- und Relativitätstheorie. Daraus ergibt sich eine wesentliche Einschränkung für die Anwendbarkeit anschaulicher Vorstellungen, die der Makrowelt angehören. Den Elementarteilchen dürfen daher nur solche Eigenschaften zugeordnet werden, für deren Existenz ein direkter experimenteller Beleg vorliegt, nicht solche, die aus hypothetischen Strukturvorstellungen folgen. Nur bei einigen wenigen Teilchen kann man begründet von einer räumlichen Ausdehnung sprechen. Im Einklang mit den Prinzipien der Quantenphysik sind die Teilchen ein und derselben Sorte ununterscheidbar (↑ 2.5.5.3.). Darüber hinaus fordert die relativistische Quantenphysik, daß zu jeder Sorte ein *Antiteilchen* auftritt. Die Umwandlungsprozesse zwischen Elementarteilchen tragen wie alle Quantenprozesse statistischen Charakter.
Die Physik der Elementarteilchen ist ein nicht abgeschlossenes Gebiet. Als gesicherte Tatsachen können zuerst einmal nur die durch wiederholte, unabhängige Experimente aufgefundenen Eigenschaften der schon bekannten Teilchen angesehen werden, dazu in gewissem Maße auch einige allgemeine theoretische Gesetzmäßigkeiten. Die speziellen Vorstellungen von der Natur der Elementarteilchen, und damit auch jede konkrete Theorie, sind dagegen offen gegenüber dem Zustrom neuer experimenteller Fakten, z. B. Entdeckung neuer Teilchen und neuer Wechselwirkungen.

3.2.1.2. Nachweis von Elementarteilchen

3.2.1.2.1. *Nachweismethoden.* Nur die wenigsten Teilchen können unmittelbar in den Nachweisgeräten beobachtet werden. Meist lassen sich nur die Zerfallsprodukte nachweisen und quantitativ erfassen, um aus ihnen dann einen Rückschluß auf das ursprüngliche Teilchen zu ziehen. Es können entweder die elektrisch geladenen langlebigen oder stabilen Teilchen p, $e^\pm$, $\mu^\pm$, $\pi^\pm$, $K^\pm$ oder Photonen nachgewiesen werden. Bei geladenen Teilchen wird der Umstand ausgenutzt, daß sie beim Durchgang durch Materie entweder eine Spur angeregter oder ionisierter Zentren (Atome oder Moleküle) längs ihrer Flugbahn hinterlassen oder zur Ausstrahlung von Licht führen. Im Falle von Photonen macht man sich den Effekt der Konversion zunutze: Hochenergetische Photonen gehen im elektrischen Feld schwerer Atome in Elektron-Positron-Paare über, die dann wie andere geladene Teilchen analysiert werden. Das Ziel der Messung ist sowohl die Identifikation des Teilchens als auch die Messung seines Impulses. Zu diesem Zweck setzt man die verschiedenen Detektoren in ein kon-

stantes Magnetfeld und bestimmt aus der Bahnkrümmung den Impuls der Teilchen (s. a. 2.1.6.6.). Die Identifikation des Teilchens läuft auf die Ermittlung seiner Masse hinaus. Dazu muß zusätzlich zum Impuls entweder die Energie oder die Geschwindigkeit bestimmt werden. Die Geschwindigkeitsbestimmung erfolgt z. B. durch Ausnutzung des Čerenkoveffekts (↑ 3.2.1.2.2.).

3.2.1.2.2. *Nachweisgeräte.* Die *Blasenkammer* ist ein Gefäß, das eine überhitzte Flüssigkeit, z. B. Wasserstoff, Deuterium oder Propan, enthält. Das hindurchfliegende Teilchen hinterläßt auf seiner Flugbahn ionisierte Atome, die den Siedevorgang einleiten. Die entstehenden Gasbläschen kennzeichnen die Spur des Teilchens, die fotografiert wird, um sie später zu analysieren (Tafeln 8/9). Die Blasenkammer hat die noch früher entwickelte *Nebelkammer* ersetzt, deren Wirkungsweise auf dem Prinzip des übersättigten Dampfs beruht. Bei *Streamerkammern* werden die sekundären Ionisationseffekte bei Vorhandensein eines kurzzeitig, d. h. etwa 10^{-9} s lang, angelegten, sehr starken elektrischen Felds ausgenutzt, durch das die primär erzeugten Elektronen und Ionen beschleunigt werden. Es kommt längs der Teilchenbahn zu bandförmigen Leuchterscheinungen, die entweder direkt oder unter Zuhilfenahme von Lichtverstärkern fotografiert werden. Eine andere Klasse von Detektoren, die *Ionisations-*, *Proportional-* und *Driftkammern*, kommt ohne eine fotografische Aufzeichnung der Teilchenspur aus. Auch hier erzeugt das nachzuweisende Teilchen im Gas der Kammer Elektronen und Ionen. Ein elektrisches Feld sorgt für Trennung und Beschleunigung dieser Ladungsträger, so daß durch sekundäre Ionisation eventuell weitere erzeugt werden. Die entstandenen Elektronen werden durch ein System paralleler Drähte aufgefangen. Diese Drähte sind so in Ebenen (z. B. parallel zur x-y-Ebene eines rechtwinkligen Koordinatensystems) gespannt, daß immer abwechselnd die Drähte einer Ebene längs der x-Achse, und die der folgenden Ebene längs der y-Achse verlaufen. Eine Registrierung des elektrischen Impulses erfolgt durch denjenigen Draht, welcher der Teilchenspur am nächsten ist. Ein einzelnes Paar solcher dicht hintereinander aufgestellter Drahtebenen fixiert also den Schnittpunkt der Spur, weitere Paare ermöglichen die Bestimmung des Kurvenverlaufs. Die gesamte in Form von elektrischen Impulsen anfallende Information wird unmittelbar in einem Computer verarbeitet.

Die Bestimmung der Geschwindigkeit ist mit Hilfe von *Čerenkovzählern* möglich, deren Wirkungsweise auf dem *Čerenkoveffekt* beruht: Elektrisch geladene Teilchen strahlen bei Durchgang durch ein Medium mit dem Brechungsindex n Licht aus, falls die Bedingung $v/c > 1/n$ erfüllt ist. Durch eine Staffelung von Medien mit unterschiedlichem Brechungsindex oder eine Messung des Abstrahlungswinkels der Čerenkovstrahlung läßt sich ein Rückschluß auf die Geschwindigkeit des hindurchfliegenden Teilchens ziehen.

Neben diesen Geräten, die der Messung eines einzelnen Teilchens dienen, gibt es sogenannte *Kalorimeter* zur Bestimmung der Energie, die ein manchmal aus vielen Teilchen bestehendes Teilchensystem hat. Die primären Teilchen übertragen ihre Energie auf einen Schauer geladener Teilchen, die ihrerseits ein Szintillatormaterial zur Lichtaussendung anregen. Die Lichtausbeute ist ein Maß für die Energie des ursprünglichen Teilchensystems. Auf diese Weise läßt sich die Energie eines Neutrinos bestimmen, das in einer Reaktion erzeugt worden ist. Sie ergibt sich als Differenz aus der Anfangsenergie der Reaktionspartner (etwa Myon und Proton) und der im Kalorimeter gemessenen Energie aller gleichzeitig mit dem Neutrino erzeugten Teilchen. Bei den neueren Experimenten der Elementarteilchenphysik werden sehr komplexe Apparaturen eingesetzt, die aus einer Kombination vieler einzelner Detektoren bestehen (Tafeln 14/15).

3.2.1.3. Meßbare Eigenschaften

Eine Zusammenstellung der wichtigsten Daten ist in Tab. 3.2.2-1 zu finden.

3.2.1.3.1. *Masse.* Genau wie ein makroskopischer Körper haben die Teilchen eine Masse. Die für das Teilchen charakteristische Ruhmasse (↑ 2.1.3.5.) wird gewöhnlich in der Einheit GeV/c^2 angegeben, wobei die Einsteinsche Beziehung $E = mc^2$ ausgenutzt wird. Das Massenspektrum der Elementarteilchen kann mit den gegenwärtig existierenden Beschleunigern bis zu einigen GeV/c^2 untersucht werden (1 GeV/c^2 = $1{,}782 \cdot 10^{-24}$ kg).

3.2.1.3.2. *Spin.* Der Eigendrehimpuls eines Teilchens heißt Spin (↑ 2.1.3.9.). Er wird in Einheiten der Planckschen Konstante $\hbar = h/2\pi$ gezählt. Im Einklang mit der Quantentheorie nimmt der Spin J nur halb- oder ganzzahlige Werte an (s. a. 2.5.3.). Masse und Spin charakterisieren immer eine Teilchensorte, etwa das Proton. Ein konkret existierendes Proton wird erst durch die zusätzliche Angabe von Impuls $\boldsymbol{p}$ und Spinkomponente J_z in Richtung einer willkürlich gewählten z-Achse vollständig bestimmt. Dabei kann $J_z/\hbar$ die $2J + 1$ Werte $-J, -J + 1, \ldots, J$ annehmen. Diese Regel gilt nicht für masselose Teilchen. So ist für die wahrscheinlich masselosen Neutrinos ν_e, $\bar{\nu}_e$, ν_μ, $\bar{\nu}_\mu$ der Spin immer entweder parallel oder antiparallel zum Impuls orientiert, so daß ν_e und ν_μ durch die linke Hand, $\bar{\nu}_e$ und $\bar{\nu}_\mu$ durch die rechte Hand veranschaulicht werden (Abb. 3.2.1-1). Diese Vorstellungen von der Händigkeit der

Neutrinos müßten korrigiert werden, falls exaktere Messungen für die Neutrinomassen zwar kleine, jedoch von Null verschiedene Werte liefern sollten. Auch die Photonen, für die $J = 1$ gilt, spielen eine Sonderrolle. Nur die Spinprojektion in Richtung des Impulses ist meßbar, sie nimmt nur die zwei Werte $+1$ oder -1 an. Der Spin kann für langlebige Teilchen aus einem Stern-Gerlach-Experiment (↑ 1.1.2.5.) bestimmt werden; im allgemeinen erschließt man ihn aus einer Analyse der Winkelabhängigkeit der Zerfallsprozesse des betreffenden Teilchens.

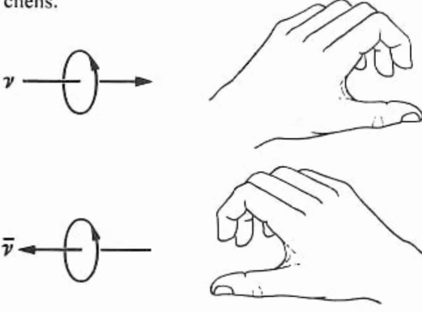


Abb. 3.2.1-1 Orientierung von Spin und Impuls (*Pfeilrichtung*) bei Neutrino und Antineutrino

3.2.1.3.3. *Lebensdauer.* Für relativ langlebige Teilchen, die in den Nachweisapparaturen noch eine beobachtbare Spur hinterlassen, kann die Lebensdauer τ direkt aus der Länge der Spur ermittelt werden. Für sehr kurzlebige Teilchen, die sogenannten Resonanzen (↑ 3.2.2.2.), ist das unmöglich. Da sich die Resonanzen als Maxima im Wirkungsquerschnitt bemerkbar machen (Abb. 3.2.1-2), kann aus der energetischen Breite Γ des Maximums gemäß der quantentheoretischen Beziehung $\Gamma\tau \approx \hbar \approx 6{,}6 \cdot 10^{-22}$ MeVs auf die mittlere Lebensdauer geschlossen werden (s. a. 2.5.6.3.). Für die Resonanz Δ (1236), die eine Breite von 120 MeV hat, findet man auf diese Weise $\tau \approx 5{,}5 \cdot 10^{-24}$ s.

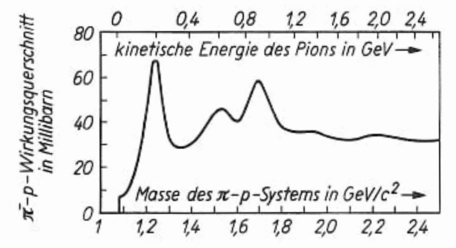


Abb. 3.2.1-2 Totaler Wirkungsquerschnitt der Streuung von negativen Pionen an Protonen

3.2.1.3.4. *Elektrische Ladung.* Die elektrische Ladung aller bisher gefundenen Teilchen ist ein positives oder negatives Vielfaches der Elementarladung e (↑ 1.1.2.1.). In dieser Einheit hat das Elektron die Ladung $Q = -1$, das Proton $Q = +1$ und das Neutron $Q = 0$. Im Falle von Teilchen, die unterschiedlich geladen auftreten können, wird die Ladung rechts oben am Symbol angegeben, z. B. K^+, K^0. Die hypothetischen Quarkteilchen (s. a. 3.2.2.2.) sollten die Ladungen $Q = \pm\frac{1}{3}$ oder $Q = \pm\frac{2}{3}$ tragen.

3.2.1.3.5. *Magnetisches Moment.* Teilchen mit Spin haben im allgemeinen ein magnetisches Moment, d. h. sie verhalten sich in einem Magnetfeld wie eine kleine Magnetnadel. Sofern diese Teilchen frei fliegend zu beobachten sind, läßt sich ihr magnetisches Moment durch eine geeignete Anordnung von Magnetfeldern experimentell bestimmen. Das magnetische Moment μ ist die Summe aus einem sogenannten normalen Anteil $e\hbar/2M$, das jedes elektrisch geladene Teilchen der Masse M nach der relativistischen Quantenmechanik besitzt, und einem anormalen Anteil, der Resultat der komplizierten Wechselwirkungen der Teilchen untereinander ist. Präzisionsmessungen von μ ergaben für Elektronen eine Abweichung von 1,1596 ‰, für Myonen von 1,1659 ‰ vom normalen Wert. Das befindet sich im Einklang mit den in der Quantenelektrodynamik errechneten Korrekturen. Die Momente von Proton, Neutron und anderen Baryonen (↑ 3.2.2.2.) weichen viel stärker von den normalen Werten ab: $\mu_p = 2{,}793 e\hbar/2M_p$, $\mu_n = -1{,}913 e\hbar/2M_p$. Darüber hinaus ist μ bisher nur für die Baryonen Λ, Σ^+, Σ^-, Ξ^0 und Ξ^- gemessen worden.

3.2.1.3.6. *Ausdehnung.* Die Ausdehnung von Teilchen kann durch die räumliche Verteilung ihrer elektrischen Ladung definiert werden. Die experimentelle Bestimmung dieser Ladungsdichte erfolgt durch elastische Streuung von Elektronen (↑ 3.2.5.1.). Für Proton und Neutron ergeben sich Radien in der Größenordnung von 1 fm = 10^{-13} cm (Abb. 3.2.1-3).

3.2.1.4. Ladungsartige Quantenzahlen. Die Charakterisierung von Teilchen erfordert über die bisher angeführten Eigenschaften hinaus die Angabe gewisser Quantenzahlen, die keine Rolle in der Makrophysik spielen und daher unanschaulich sind. Sie sind für das Einzelteilchen nicht direkt meßbar, sondern dienen zur Klassifizierung der Teilchen und zur Formulierung von Einschränkungen bei Umwandlungsprozessen.

3.2.1.4.1. *Baryonzahl.* Den Baryonen (↑ 3.2.2.2.) wird die Baryonzahl $B = 1$ zugeordnet, ihren Antiteilchen $B = -1$ und allen übrigen Teilchen $B = 0$. Demnach hat das Proton $B = +1$, Elektron, Photon und Pion haben $B = 0$. B wird als eine Quantenzahl eingeführt, für welche ein absoluter Erhaltungssatz (↑ 3.2.4.1.)

gilt. Dadurch werden unter anderem Reaktionen von der Art $p + p \rightarrow \pi^+ + \pi^+$ oder $p + e^- \rightarrow 2\gamma$ ausgeschlossen und damit die Stabilität der massiven Materie theoretisch erfaßt. Allerdings geben die Vorstellungen von einer möglichen Vereinigung der schwachen, elektromagnetischen und starken Wechselwirkungen (s. a. 3.2.6.3.) Veranlassung, diesen Erhaltungssatz einer experimentellen Prüfung zu unterziehen. Es wird insbesondere nach Zerfällen des Protons gesucht, um eine eventuelle schwache Verletzung der Erhaltung von B quantitativ zu bestimmen.

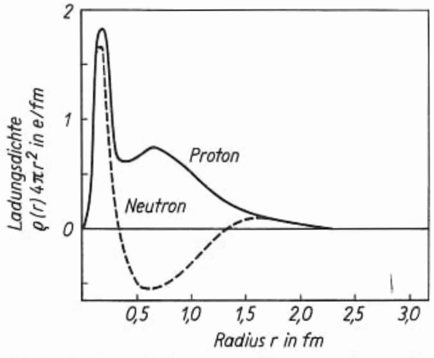


Abb. 3.2.1-3 Ladungsverteilungen der Nukleonen

3.2.1.4.2. *Leptonzahlen.* Jedem Paar von Leptonen (↑ 3.2.2.1.) ist eine eigene Leptonzahl zugeordnet. So haben das Elektron e^- und sein Neutrino ν_e die Elektronenzahl $L_e = +1$, ihre Antiteilchen e^+ und $\bar{\nu}_e$ aber $L_e = -1$. Entsprechend ordnet man den anderen Leptonen μ^-, τ^- Leptonzahlen zu. Der Erhaltungssatz für L_e oder L_μ verbietet beispielsweise den Zerfall $\mu^- \rightarrow e^- + \gamma$ oder den Prozeß $\bar{\nu}_e + n \rightarrow p + e^-$.

3.2.1.4.3. *Hyperladung.* Die Hyperladung Y dient der weiteren Unterscheidung von Hadronen (↑ 3.2.2.2.). Die möglichen Werte von Y sind ganzzahlig. Die Einführung der Hyperladung hat sich als zweckmäßig erwiesen, um die Gesetzmäßigkeiten bei der Erzeugung der K-Mesonen und der Baryonen Σ, Λ, Ξ zu erfassen (s. a. 3.2.2.2.).

3.2.1.4.4. *Charm.* Die Entdeckung der neuen Teilchen ψ, D, ... hat die Einführung der Quantenzahl Charm C nahegelegt, die gleichfalls nur für Hadronen definiert ist. Allen vorher bekannten Hadronen wird der Wert $C = 0$ zugeordnet. Im Rahmen des Quarkmodells (↑ 3.2.3.2.) bedeutet dies, ein viertes Quarkteilchen mit $C = +1$ einzuführen, während gleichzeitig die ursprünglichen Quarkteilchen die Quantenzahl $C = 0$ erhalten. Charm und Hyperladung spielen eine weitgehend ähnliche Rolle als Ladungen, die nur in der Mikrowelt von Bedeutung sind. Die kürzliche Entdeckung des schweren Mesons Y (9,4 GeV/c^2) macht vermutlich die Erfindung einer weiteren Quantenzahl erforderlich.

3.2.1.4.5. *Strangeness.* Gelegentlich benutzt man auch die Quantenzahl Strangeness S [engl., Seltsamkeit], die mit B, Y und C durch die Beziehung $Y = B + S - C$ verbunden ist.

3.2.1.5. Isospin. Im Massenspektrum der Hadronen zeigen sich Multipletts von Teilchen mit relativ kleinen Massenunterschieden, die sich durch ihre elektrische Ladung voneinander unterscheiden, aber den gleichen Spin und übereinstimmende Quantenzahlen B, Y, C besitzen. Beispiele dafür sind die Dubletts (p, n) und (K^+, K^0), die Tripletts (Σ^-, Σ^+, Σ^+) und das Singulett Λ, das keine Multiplettpartner hat. Es liegt nahe und ist am Beispiel des Nukleonendubletts zuerst durchgeführt worden, die zu einem Multiplett gehörenden Teilchen als verschieden geladene Varianten eines Teilchens aufzufassen. Der mathematischen Durchführung dieser Konzeption dient der Formalismus des Isospins, der analog zur Quantentheorie des Drehimpulses (↑ 2.5.3.) aufgebaut wird: Ein in einem abstrakten Raum definierter Isospinvektor mit dem Quadrat $I(I + 1)$ hat $2I + 1$ verschiedene Einstellmöglichkeiten, wobei eine Komponente $I_3 = -I, -I + 1, \ldots + I$ insgesamt $2I + 1$ verschiedene Werte durchläuft. Benachbarten Werten von I_3 entsprechen Teilchen, deren elektrische Ladungen Q sich um 1 unterscheiden. Q ist durch I_3, die Hyperladung Y und die Charmquantenzahl C bestimmt: $Q = I_3 + \frac{1}{2}(Y + 2C)$. Ebenso wie der Spin J nimmt auch der Isospin I nur halb- oder ganzzahlige Werte an; alle Teilchen eines Multipletts haben gleichen Isospin. $I = 0$ liefert ein Singulett, $I = \frac{1}{2}$ ein Dublett, $I = 1$ ein Triplett usw.

3.2.1.6. Parität. Allen Teilchen mit Ausnahme des Neutrinos ist eine Parität P zugeordnet, die

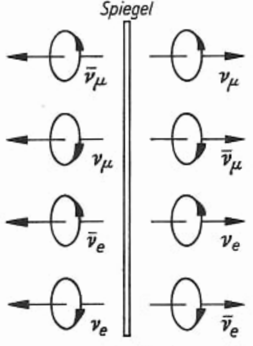


Abb. 3.2.1-4 Anwendung der Spiegelungstransformation auf Neutrinos

die Werte $+1$ oder -1 annimmt. Bei einer *Raumspiegelung* geht z. B. die quantentheoretische Wellenfunktion χ eines Zweiteilchensystems mit dem Bahndrehimpuls l bis auf ein Vorzeichen in sich über: $\chi \to (-1)^l P_1 P_2 \chi$. Auf diese Weise gehen die Paritäten P_1, P_2 der beteiligten Teilchen in das Transformationsgesetz ein (↑ 3.2.4.1.). Die Ausnahmerolle der Neutrinos erklärt sich aus ihrer Links- bzw. Rechtshändigkeit: Unter einer Spiegelung kehrt der Impuls seine Richtung um, der Spin jedoch nicht, so daß aus einem linkshändigen Neutrino ein rechtshändiges Teilchen wird, das mit dem Antineutrino identifiziert werden muß. Entsprechendes gilt für das Antineutrino (Abb. 3.2.1-4), (s. a. 3.2.1.3.2.). Die Paritäten lassen sich aus der Winkelabhängigkeit der Wirkungsquerschnitte bestimmen.

3.2.1.7. Antiteilchen. Zu jedem Teilchen gibt es ein Antiteilchen, dessen Quantenzahlen in einer typischen Beziehung zu denen des Teilchens stehen (Tab. 3.2.1-5). Das Antiteilchen wird durch einen Querstrich über dem Symbol bezeichnet, z. B. $\bar{p}$. Abweichend von dieser Regel bezeichnet e^+ das Antielektron oder Positron, μ^+ das Antimyon. Falls alle ladungsartigen Quantenzahlen eines Teilchens den Wert Null haben, ist dieses mit seinem Antiteilchen identisch, z. B. $\bar{\pi}^0 = \pi^0$, $\bar{\gamma} = \gamma$. Die Antiteilchen stimmen in Masse und Lebensdauer mit ihrem Partner überein. Insbesondere sind $\bar{p}$ und e^+ stabile Teilchen. Ihre scheinbare Instabilität im Laboratorium rührt daher, daß sie in einer Umgebung, die Protonen und Elektronen enthält, mit großer Wahrscheinlichkeit einer Paarvernichtung (s. u.) unterliegen. Es ist von vornherein nicht auszuschließen, daß in anderen Teilen des Universums *Antimaterie* in makroskopischem Ausmaß existiert, die aus Antikernen, Antiatomen, Antimolekülen usw. besteht. Eine für die Teilchenphysik charakteristische Besonderheit besteht darin, daß zu jedem Prozeß ein anderer möglicher Prozeß gehört, in dem ein oder mehrere Teilchen aus dem Anfangs- in den Endzustand oder umgekehrt gewechselt haben und nun als Antiteilchen erscheinen; z. B. ist mit $a + b \to c + d$ auch $a + \bar{c} \to \bar{b} + d$ ein möglicher Prozeß.

Ein Blick auf Tab. 3.2.1-5 zeigt, daß jedes aus Teilchen und Antiteilchen bestehende System die Quantenzahlen $B = L_e = L_\mu = 0$, $I = Q = Y = C = 0$ hat, und daher ungehindert durch Erhaltungssätze (↑ 3.2.3.1.) in jedes andere aus einem Teilchen und seinem Antiteilchen bestehende System übergehen kann. Es sind z. B. die folgenden Prozesse möglich: $e^- + e^+ \to 2\gamma$, $e^+ + e^- \to p + \bar{p}$, $e^- + e^+ \to \tau^- + \tau^+$. Man spricht hier von einer *Paarvernichtung* im Anfangs- bzw. einer *Paarerzeugung* im Endzustand. Solche Prozesse werden an den Elektron-Positron-Speicherringen untersucht (s. a. 3.2.5.2.). Bei der Erzeugung schwerer Teilchen wandelt sich die kinetische Energie von Elektron und Positron in Ruhmasse der schweren Teilchen um. Der umgekehrte Effekt liegt bei der Zerstrahlung eines massiven Paars in Photonen vor. Die Mindestenergie, die für die Erzeugung schwerer Teilchen erforderlich ist, bezeichnet man als *Schwellenergie*. Bei dieser Energie werden die neuen Teilchen mit verschwindend kleinem Impuls erzeugt, so daß die Schwellenergie gleich der doppelten Ruhenergie des zu erzeugenden Teilchens ist.

Tab. 3.2.1-5 Relationen zwischen den Quantenzahlen eines Teilchens und des zugehörigen Antiteilchens

Teilchen	Antiteilchen
B	$-B$
L_e	$-L_e$
L_μ	$-L_\mu$
Q	$-Q$
I_3	$-I_3$
Y	$-Y$
I	I
J	J
P	$(-1)^{2J} P$

3.2.2. Klassifikation der Elementarteilchen

Eine natürliche Einteilung der Teilchen beruht darauf, daß sie in unterschiedlichem Maße den verschiedenen Typen von Wechselwirkungen (↑ 3.2.4.) unterliegen. Es ergibt sich eine Einteilung in Leptonen, Hadronen und die intermediären Bosonen, die ein Bindeglied zwischen Leptonen und Hadronen darstellen, indem sie an beide ankoppeln. Zur Zeit ist allerdings nur ein einziger Vertreter dieser Gruppe nachgewiesen, nämlich das Photon. Teilchen und Antiteilchen gehören jeweils ein und derselben Familie an. In Tab. 3.2.2-1 sind die relativ langlebigen Teilchen mit ihren Quantenzahlen zusammengestellt.

3.2.2.1. Leptonen. In diese Gruppe gehören das Elektron e^- und das Elektron-Neutrino ν_e, das Myon μ^- und das Myon-Neutrino ν_μ sowie das schwere Lepton τ. Der Umstand, daß Elektron und Myon sehr leichte Teilchen ($M_e = 0{,}511$ MeV$/c^2$, $M_\mu = 106$ MeV$/c^2$) und die Neutrinos wahrscheinlich masselos sind, hat zu der Bezeichnung Leptonen, d. h. leichte Teilchen, Anlaß gegeben. Die Eigenart der Leptonen besteht darin, nicht der starken Wechselwirkung zu unterliegen (↑ 3.2.4.3.). Das Myon zerfällt im Prozeß $\mu^- \to e^- + \bar{\nu}_e + \nu_\mu$ mit einer Lebensdauer von $2{,}2 \cdot 10^{-6}$ s. Das Neutrino ist ursprünglich nur hypothetisch eingeführt worden, um die scheinbare Verletzung der Energieerhaltung beim β-Zerfall $n \to p + e^- + \bar{\nu}_e$ zu

erklären (↑ 3.4.1.2.1.). Da das Neutrino ungeladen ist, unterliegt es ausschließlich der schwachen Wechselwirkung mit dementsprechend kleinen Wirkungsquerschnitten, und kann daher nur äußerst schwer nachgewiesen werden.

Der *Nachweis der Neutrinos* erfolgt in der Regel aufgrund der durch sie initiierten Reaktionen (Tafel 13). Das sind in der Nachweisreaktion der Elektron-Neutrinos ν_e Kernumwandlungen, z. B. $\nu_e + {}^{37}\text{Cl} \rightarrow {}^{37}\text{Ar} + e^-$, deren Endprodukt quantitativ erfaßt wird. Antineutrinos $\bar{\nu}_e$ lassen sich mit Hilfe der Reaktion $\bar{\nu}_e + p \rightarrow n + e^+$ nachweisen. Beobachtet werden in diesem Falle die Photonen, die bei den nachfolgenden Einfangprozessen von n und e^+ entstehen. Myon-Neutrinos werden durch die in den Reaktionen ν_μ + Nukleon $\rightarrow \mu^- + \ldots$ und $\bar{\nu}_\mu$ + Nukleon $\rightarrow \mu^+ + \ldots$ gebildeten μ^- und μ^+ angezeigt. Der experimentelle Beleg für die Verschiedenheit von ν_e und ν_μ besteht darin, daß die beim Pionzerfall entstandenen Neutrinos bei ihrer Reaktion mit Nukleonen immer nur Myonen, jedoch niemals Elektronen erzeugen; es handelt sich also um Myon-Neutrinos. Wenn sich die Neutrinomassen als von Null verschieden erweisen sollten, könnten sich auch die Eigenschaften der Neutrinos als verwickelter herausstellen, als bisher angenommen worden ist. Es bestände dann die Möglichkeit, daß man unterscheiden müßte zwischen den physikalischen, d. h. als freie Teilchen existie-

Tab. 3.2.2-1 Tabelle der wichtigsten Teilchen; nd nicht definiert

	M in GeV/c²	τ in s	Q in |e|	J	P	I	I_3	Y	C	B	L_e	L_μ	wichtige Zerfälle
Photon γ	0	∞	0	1	−1	nd	nd	0	0	0	0	0	
Leptonen													
Elektron e^-	$5{,}110 \cdot 10^{-4}$	∞	−1	1/2	+1	nd	nd	nd	nd	0	+1	0	
Myon μ^-	0,1057	$2{,}20 \cdot 10^{-6}$	−1	1/2	+1	nd	nd	nd	nd	0	0	+1	$e^-\bar{\nu}\nu_{e\mu}$
Schweres Lepton τ^-	1,782	?	−1	1/2	1/2	nd	nd	nd	nd	0			$\mu^-\bar{\nu}\nu$, $e^-\bar{\nu}\nu$
Elektron-Neutrino ν_e	0	∞	0	1/2	nd	nd	nd	nd	nd	0	+1	0	
Myon-Neutrino ν_μ	0	∞	0	1/2	nd	nd	nd	nd	nd	0	0	+1	
Mesonen													
Pion π^0	0,1350	$0{,}83 \cdot 10^{-16}$	0	0	−1	1	0	0	0	0	0	0	2γ
Pion π^+	0,1396	$2{,}6 \cdot 10^{-8}$	+1	0	−1	1	+1	0	0	0	0	0	$\mu^+\nu_\mu$
Eta-Meson η	0,549	$\Gamma = 0{,}8$ keV	0	0	−1	0	0	0	0	0	0	0	2γ, $3\pi^0$
Kaon K^+	0,494	$1{,}24 \cdot 10^{-8}$	+1	0	−1	1/2	+1/2	1	0	0	0	0	$\mu^+\nu_\mu$, $\pi^+\pi^0$, $2\pi^+\pi^-$
Kaon K^0	0,498	50% $8{,}9 \cdot 10^{-11}$	0	0	−1	1/2	−1/2	1	0	0	0	0	$\pi^+\pi^-$, $2\pi^0$
$\bar{K}^0$		50% $5{,}2 \cdot 10^{-8}$				1/2	+1/2	−1					$3\pi^0$, $\pi^+\pi^-\pi^0$
D-Meson D^+	1,868		1	0	−1				1	0	0	0	$K^-\pi^+\pi^+$
D-Meson D^0	1,863		0	0	−1				1	0	0	0	$K^-\pi^+\pi^0$
F-Meson F^+	2,03		1	0	−1		0		1	0	0		
Baryonen													
Proton p	0,9383	∞	+1	1/2	+1	1/2	+1/2	1	0	1	0	0	$\bar{p}$
Neutron n	0,9396	918	0	1/2	+1	1/2	−1/2	1	0	1	0	0	$pe^-\nu_e$
Lambda Λ	1,116	$2{,}6 \cdot 10^{-10}$	0	1/2	+1	0	0	0	0	1	0	0	$p\pi^-$, $n\pi^0$
Sigma Σ^+	1,189	$0{,}8 \cdot 10^{-10}$	+1	1/2	+1	1	+1	0	0	1	0	0	$p\pi^0$, $n\pi^+$
Sigma Σ^0	1,192	$5{,}8 \cdot 10^{-20}$	0	1/2	+1	1	0	0	0	1	0	0	$\Lambda\gamma$
Sigma Σ^-	1,197	$1{,}4 \cdot 10^{-10}$	−1	1/2	+1	1	−1	0	0	1	0	0	$n\pi^-$
Xi θ^0	1,315	$2{,}9 \cdot 10^{-10}$	0	1/2	+1	1/2	−1/2	−1	0	1	0	0	$\Lambda\pi^0$
Xi θ^-	1,321	$1{,}6 \cdot 10^{-10}$	−1	1/2	+1	1/2	+1/2	−1	0	1	0	0	$\Lambda\pi^-$
Omega Ω^-	1,672	$\sim 1 \cdot 10^{-10}$	−1	3/2	+1	0	0	−2	0	1	0	0	$\theta^0\pi^-$, $\theta^-\pi^0$, ΛK^-
Lambda-Charm Λ_c^+	2,295	$7 \cdot 10^{-13}$	+1	1/2					1	1	0	0	$p\pi^+K^-$
Sigma-Charm Σ_c^{++}	2,426		+2	1/2					1	1	0	0	$\Lambda\pi^+\pi^+\pi^+\pi^0$

renden Neutrinos, und den in den Wechselwirkungsprozessen auftretenden ν_e, $\bar{\nu}_e$, ν_μ und $\bar{\nu}_\mu$. Die Erhaltungssätze für L_e und L_μ hätten nur für die Wechselwirkungsakte Geltung, wogegen die freien Neutrinos als quantentheoretische Überlagerung aus ν_e, ν_μ, $\bar{\nu}_e$ und $\bar{\nu}_\mu$ aufzufassen wären. Nach diesen sogenannten *Neutrinooszillationen* wird gegenwärtig intensiv gesucht.

Heute ist es möglich, mit allen Arten von Neutrinos Streuexperimente durchzuführen (↑ 3.2.5.3.). Diese Neutrinos stammen entweder aus dem β-Zerfall der in Kernreaktoren erzeugten Neutronen oder aus dem Zerfall von Kaonen oder Pionen gemäß $K^- \to \mu^- + \bar{\nu}_\mu$ oder $K^+ \to \mu^+ + \nu_\mu$.

Im Unterschied zu den meisten Elementarteilchen ist das Elektron-Neutrino eine Komponente der natürlichen irdischen Umwelt. Die im Innern der Sonne und anderer Sterne ablaufenden Kernreaktionen sind Quelle eines Stroms von Neutrinos ν_e, die praktisch ungeschwächt sowohl die Atmosphäre als auch den Erdkörper durchdringen. Gegenwärtig werden große Anstrengungen unternommen, um diesen Teilchenstrom quantitativ zu erfassen. Ein solches Großexperiment wird im Kaukasus (UdSSR) vorbereitet, mehr als 1500 m unter Tage, um jegliche Störung durch Teilchen der Höhenstrahlung auszuschließen. Die Menge des in riesigen Kavernen exponierten Detektormaterials beträgt einige Tausend Tonnen. Die Stabilität und die kleinen Wirkungsquerschnitte führen zu dem Schluß, daß selbst in kosmischen Dimensionen Neutrinos kaum absorbiert werden. Es sollte daher eine universelle Neutrino-Untergrundstrahlung existieren.

Eine der bedeutungsvollsten Entdeckungen der letzten Jahre ist die des *schweren Leptons* τ^- und seines Antiteilchens τ^+. Die Familie der Leptonen hatte seit dreißig Jahren für abgeschlossen gegolten. Gesichert ist, daß τ ein einfach geladenes Teilchen mit dem Spin $J = 1/2$ und einer Masse von etwa 1,8 GeV/c^2 ist. Das neue Lepton wurde an den Elektron-Positron-Speicherringen dadurch entdeckt, daß Endzustände beobachtet wurden, die nur aus μ^-, e^+ oder μ^+, e^- und zusätzlich aus Neutrinos bestehen. Man erklärt sich dies so, daß zuerst ein τ-Paar in der Reaktion $e^+ + e^- \to \tau^+ + \tau^-$ erzeugt worden ist, dessen Partner anschließend in den Reaktionen $\tau^+ \to e^+ + \nu_e + \bar{\nu}$ und $\tau^- \to \mu^- + \bar{\nu}_\mu + \nu$ zerfallen. Es ist noch völlig ungeklärt, ob hier auch ein besonderes τ-Neutrino in Erscheinung tritt, und wenn ja, ob es masselos ist. Der Umstand, daß τ offenbar nicht der starken Wechselwirkung unterliegt, rechtfertigt die Bezeichnung Lepton, obwohl seine Masse groß ist. In einigen der Zerfallskanäle von τ^- konnte man erstmalig Hadronen als Zerfallsprodukte eines Leptons beobachteten, z. B. das Pion in dem Zerfall $\tau^- \to \pi^- + \nu$.

3.2.2.2. Hadronen. Die charakteristische Eigenschaft der Hadronen ist es, allen Typen von Wechselwirkungen zu unterliegen. Die Hadronen mit halbzahligem Spin heißen *Baryonen*, diejenigen mit ganzzahligem Spin *Mesonen*. Sehr kurzlebige Hadronen, wie etwa die zur Zeit bekannten Mesonen mit dem Spin 1 und höher werden *Resonanzen* genannt. Sie zeigen sich nur als Maxima in der Energieabhängigkeit der Wirkungsquerschnitte. Die bekanntesten Baryonen sind Proton und Neutron, die auch als *Nukleonen* bezeichnet werden, da sie die Bausteine der Atomkerne sind. Das *Proton* ist stabil, das *Neutron* mit einer mittleren Lebensdauer von 918 s quasistabil. Wenn das Neutron sich jedoch in stabilen Atomkernen oder in Neutronensternen (↑ 10.3.3.9.2.) befindet, kann der Zerfall energetisch unmöglich sein. Das einfachste Beispiel bietet der Deuteriumkern, der ein gebundenes System aus Proton und Neutron ist. Wenn das Neutron durch β-Zerfall (↑ 3.3.1.2.) in ein Proton übergehen wollte, müßte es zugleich den Kern verlassen, da es keinen aus zwei Protonen bestehenden Kern gibt. Dafür aber reicht die ursprüngliche Energie nicht aus.

Alle Hadronen außer Proton und Neutron müssen durch Stoßprozesse erzeugt werden. Es dominiert hierbei selbstverständlich die starke Wechselwirkung, die die größten Wirkungsquerschnitte ergibt. Die in den Stößen erzeugten Teilchen darf man sich nicht als Bestandteile der aufeinandertreffenden Stoßpartner vorstellen. Bei den Erzeugungsprozessen gelten Gesetzmäßigkeiten, die man aufgrund des Erhaltungssatzes (↑ 3.2.4.1.) für die *Hyperladung* verstehen kann. Danach ist z. B. der Prozeß $p + p \to p + \Sigma^+ + K^0$ möglich, nicht aber $p + p \to \Sigma^+ + \Sigma^0 + \pi^+$. Die historische Entwicklung verlief natürlich so, daß die Quantenzahl Y eingeführt worden ist, um die experimentell beobachteten Regeln zu begründen.

Die Zerfallsprozesse der Hadronen sind im allgemeinen komplizierter als die der Leptonen (↑ Abb. 3.2.4-10). Von entscheidender Bedeutung ist, nach welchem Typ von Wechselwirkung der Zerfall abläuft. Die Benutzung des Begriffs Wechselwirkung in diesem Zusammenhang läßt sich dadurch veranschaulichen, daß der umgekehrte Prozeß ein Wechselwirkungsakt im eigentlichen Sinne ist. In der Regel ist die Lebensdauer um so kürzer, je stärker die Zerfallswechselwirkung ist. Zerfallsprozesse, die unter *Erhaltung von Isospin und Hyperladung* ablaufen, wie etwa der Zerfall der Baryonresonanz Σ^{*+} ($\Sigma^{*+} \to \Sigma^+ + \pi^0$), gehorchen der starken Wechselwirkung und beschränken die Lebensdauer auf die Größenordnung 10^{-23} s. Wenn Isospin und Hyperladung nicht erhalten

bleiben, z. B. in $\Sigma^+ \rightarrow n + \pi^+$, so erfolgt der Zerfall nach der schwachen Wechselwirkung, und die Lebensdauer ist von der Größenordnung 10^{-10} s. Ähnliches gilt für die Baryonen Σ^-, Λ, Ξ^-, Ξ^0 und die Mesonen K^+ und K^-.

Eine besonders interessante Situation liegt bei den *neutralen Kaonen* vor. Sie werden in einem Prozeß der starken Wechselwirkung als Teilchen K^0 oder $\overline{K}^0$ erzeugt, denen die Hyperladung $Y = +1$ bzw. $Y = -1$ zugeordnet ist. Nachdem die Kaonen das Gebiet verlassen haben, wo sie erzeugt worden sind, kommt es im allgemeinen zu keinem weiteren Akt der starken Wechselwirkung; allein die schwache Wechselwirkung ist maßgeblich für die weitere zeitliche Entwicklung. Als beobachtbare Teilchen existieren sie daher nur in den Zuständen

Tab. 3.2.2-2 Resonanzen im e^--e^+-Kanal

Symbol	Masse in GeV/c²	Breite in MeV/c²
ϱ^0	0,770	150
ω	0,783	10
Φ	1,020	4
ψ	3,097	0,07
ψ'	3,685	0,23
Y	9,4	
Y'	10,0	

K_S^0 und K_L^0, die als quantentheoretische Superpositionen der beiden ursprünglichen Zustände K^0 und $\overline{K}^0$ aufzufassen sind, und denen kein bestimmter Wert von Y zugeordnet werden kann. Dieser Sachverhalt läßt sich auch so ausdrücken, daß 50% aller K^0 (oder $\overline{K}^0$) als K_S^0 zerfallen, und die restlichen 50% als K_L^0. Beim Zerfall des längerlebigen Zustands K_L^0 ist der bislang einzige Prozeß entdeckt worden, der unter CP-Verletzung abläuft (↑ 3.2.4.4.).

Einen wesentlichen Zuwachs hat die Familie der Hadronen durch die sogenannten *neuen Teilchen* erfahren. Sie zeigten sich erstmalig als Resonanzen im Wirkungsquerschnitt der Elektron-Positron-Vernichtung (↑ 3.2.1.7., 3.2.5.2.). Diese Entdeckung hat darum so große Überraschung ausgelöst, weil überhaupt nicht zu erwarten war, daß im Massenbereich oberhalb von 2 GeV/c^2 noch derartig scharfe Resonanzen existieren. Die Resonanzbreite von ψ (3,097 GeV/c^2) beträgt 0,06 MeV/c^2, und auch die Breite von ψ' (3,685 GeV/c^2) ist mit 0,22 MeV/c^2 noch sehr klein. Die neuen Teilchen ψ und ψ' sind in Vielem mit den schon länger bekannten Resonanzen ϱ^0, ω und Φ zu vergleichen (Tab. 3.2.2-2). Alle treten sie im e^--e^+-Kanal auf und haben die Quantenzahlen $J = 1$ und $P = -1$. Die ungewöhnliche Schärfe von ψ und ψ' findet im Quarkmodell (↑ 3.2.3.2.) eine hinreichende Erklärung. Es ist aber nicht allein die Existenz von ψ und ψ', die Erstaunen erregt hat, sondern der Umstand, daß kurz darauf noch einige Resonanzen gefunden wurden, die in enger Beziehung zu ψ und ψ' stehen (↑ 3.2.3.3.).

In den Deutungen dieser Fakten hat sich die Auffassung durchgesetzt, daß sich hier ein neuer Freiheitsgrad der Mikrowelt manifestiert, der durch die Quantenzahl *Charm* ausgedrückt wird. Folgt man diesen Vorstellungen, so sollte es auch neue Mesonen ($J = 0$) und Baryonen ($J = 1/2$) geben. Die folgenden Mesonen sind tatsächlich gefunden worden: D^0 (1,863 GeV/c^2), D^+ (1,868 GeV/c^2) und F^+ (2,03 GeV/c^2) und ebenfalls die Baryonen Λ_c^+ (2,273 GeV/c^2) und Σ_c^{++} (2,426 GeV/c^2). Es kann als gesichert gelten, daß ihr Zerfall durch die schwache Wechselwirkung bedingt ist. Eine Bestimmung der Quantenzahlen steht jedoch noch aus. Inzwischen hat man im e^--e^+-Kanal die sehr schwere Resonanz Y (9,4 GeV/c^2) beobachtet, die möglicherweise genau wie die ψ-Resonanzen Vorbote einer Gruppe noch unbekannter Teilchen ist (↑ 3.2.3.3.).

3.2.2.3. Reggetrajektorien. Hierbei handelt es sich um eine Zusammenfassung von Teilchen und Resonanzen, die durch übereinstimmende Werte von Ladung, Baryonzahl, Isospin, Hyperladung und Parität gekennzeichnet sind. Diese Resonanzen liegen in einem M^2-J-Diagramm auf Kurven, die als Reggetrajektorien bezeichnet werden. Ihr Verlauf, soweit er durch die bekannten Resonanzen bestimmt ist,

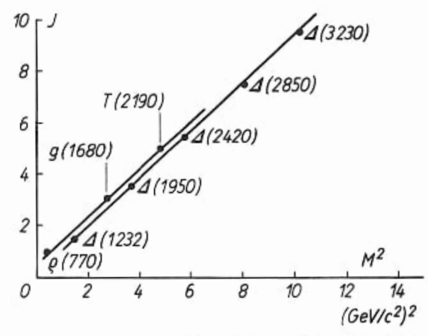


Abb. 3.2.2-3 Reggetrajektorien für die Baryonen mit den Quantenzahlen $B = 1$, $Y = 1$, $I = 3/2$, $P = 1$ und die Mesonen mit $B = 0$, $Y = 0$, $I = 1$, $P = -1$

ist annähernd linear ansteigend und parallel (Abb. 3.2.2-3). Auf der Grundlage dieses Ordnungsschemas sind in zahlreichen Fällen Resonanzen vorhergesagt worden.

3.2.3. Unitäre Symmetrie

3.2.3.1. SU(3)-Multipletts der Hadronen. Die unitären Symmetrien gestatten es, Ordnungsschemata für die Hadronen aufzustellen, die umfassender sind als die Isospinmultipletts. Physikalischer Ausgangspunkt war vor der Entdeckung der neuen Teilchen folgender Sachverhalt. Es gibt die acht Baryonen p, n, Σ^+, Σ^0, Σ^-, Λ, Ξ^0, Ξ^- mit dem Spin $J = 1/2$ und einheitlicher Parität. Sie ähneln einander hat sich bei der erfolgreichen theoretischen Vorhersage neuer Teilchen und Resonanzen erwiesen.

Das Auftreten der „magischen" Multiplizitäten 8 und 10 läßt sich aus den mathematischen Gesetzmäßigkeiten von Symmetriegruppen verstehen, und insbesondere durch die Wahl der mit dem Symbol SU(3) gekennzeichneten Symmetrie, die unter anderem Oktetts und Dekupletts, d. h. Zehnermultipletts aufweist. Die Multipletts werden durch ihre Multiplizität

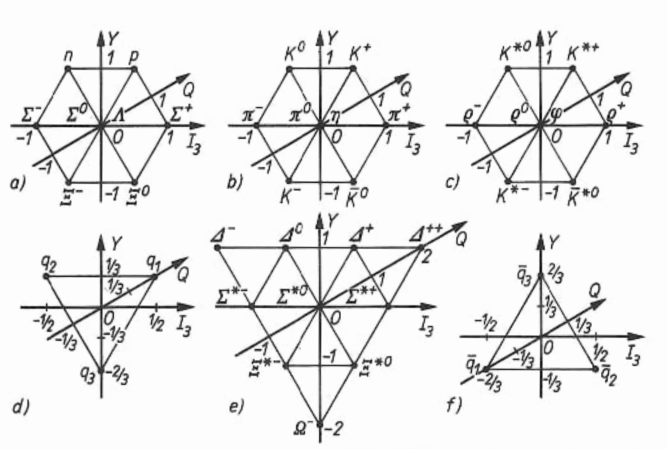


Abb. 3.2.3-2 Multipletts der Hadronen; *a)* Oktett der Baryonen mit Spin 1/2 und Parität +1, *b)* Oktett der pseudoskalaren Mesonen, *c)* Oktett der Vektormesonen, *d)* Triplett der Quarks, *e)* Dekuplett der Baryonen mit Spin 3/2 und Parität +1, *f)* Triplett der Antiquarks

darin, daß sie keine Resonanzen sind, d. h., sie sind relativ langlebig. Dasselbe gilt für die acht Mesonen π^+, π^-, π^0, η, K^+, K^0, $\bar{K}^0$, K^- mit dem Spin $J = 0$, die in demselben Maße miteinander verwandt zu sein scheinen. Solche Gruppierungen sind ebenfalls unter den Resonanzen zu beobachten, z. B. bei den zehn Baryonenresonanzen mit dem Spin $J = 3/2$ und den acht Mesonresonanzen mit $J = 1$. Wenn hierin eine Ähnlichkeit von Teilchen und damit eine Symmetrie erkannt wird, so bleiben zunächst die Unterschiede in den Massen außer Betracht. Diese Massendifferenzen sieht man als etwas Sekundäres an, nämlich als das Ergebnis einer Symmetriebrechung, die sich quantentheoretisch behandeln läßt. Die Fruchtbarkeit dieser physikalischen Auffassung

Tab. 3.2.3-1 Einfachste unitäre Symmetriegruppen; Auswahl von Multipletts mit physikalischer Bedeutung

Gruppe	Quantenzahlen	Multipletts
SU(2) (Isospin)	I_3	②, ③, ④, ...
SU(3)	I_3, Y	③, ③̄, ⑧, ⑩
SU(4)	I_3, Y, C	④, ④̄, ⑮, ⑳̄

gekennzeichnet. So bezeichnet ⑧ das Oktett, und ③ bzw. ③̄ zwei voneinander verschiedene Tripletts. Zur Unterscheidung der Mitglieder eines Multipletts ist die Angabe von zwei Quantenzahlen erforderlich, der dritten Komponente des Isospins I und der Hyperladung Y (Tab. 3.2.3-1). Jedes Oktett muß zwei Isospindubletts, ein Isospintriplett und ein Isospinsingulett enthalten (Abb. 3.2.3-2). Es zeugt von der Konsistenz des SU(3)-Schemas, daß die unabhängig davon und schon vorher erfolgte Kennzeichnung durch Isospin und Hyperladung mit der theoretisch geforderten Aufteilung in Isospinmultipletts übereinstimmt.

Physikalisch einleuchtende Vorstellungen von der Art der Symmetriebrechung führen zu den sogenannten *Massenformeln*, die experimentell gut bestätigt sind. Für das Baryonoktett lautet sie $(M_N + M_\Xi)/2 = (3M_\Lambda + M_\Sigma)/4$ und für das Baryondekuplett $M_{\Omega^-} - M_{\Xi^*} = M_{\Xi^*} - M_{\Sigma^*} = M_{\Sigma^*} - M_\Delta$. Hierbei sind immer die mittleren Massen der Isospinmultipletts gemeint. Die letzte Formel besagt, daß die einzelnen Isospinmultipletts im Dekuplett gleichweit voneinander entfernt liegen. Aus dieser Einsicht ist die Existenz des Baryons Ω^- und auch der Wert seiner Masse theoretisch vorausgesagt worden. Eine Eigentümlichkeit der SU(3)-Multipletts besteht darin, daß sie Teilchen mit in einigen Fällen sehr unterschied-

licher Lebensdauer in sich vereinen. Beispielsweise sind das stabile Proton und das kurzlebige Σ^0 in ein und demselben Oktett, ebenso die Resonanzen Δ und das langlebige Ω^- im Dekuplett. Man kann das so erklären, daß aufgrund der Erhaltungssätze (↑ 3.2.4.1.) der Zerfall der einzelnen Teilchen durch unterschiedliche Wechselwirkungen verursacht wird, was zu unterschiedlichen Lebensdauern führt. Die Klassifizierung von Hadronen in unitäre Multipletts einheitlicher Spins ist der Gruppierung von Teilchen auf Reggetrajektorien (↑ 3.2.2.3.) komplementär entgegengesetzt.

3.2.3.2. Quarks. Nach der erfolgreichen Identifikation der Hadronmultipletts bot sich die Vorstellung an, auch den zwei Tripletts eine physikalische Bedeutung zuzusprechen. Damit eröffnete sich die Möglichkeit, die Vielfalt der beobachteten Hadronen auf einige wenige fundamentale Einheiten zurückzuführen, ganz im Sinne der alten atomistischen Methode. Die drei hypothetischen Teilchen im Triplett ③ werden als die Quarks u, d, s bezeichnet, entsprechend im Triplett $\bar{③}$ die *Antiquarks* $\bar{u}$, $\bar{d}$, $\bar{s}$ (Tab. 3.2.3-3). Die Zuordnung der gebrochenen Werte von I_3, Y und Q ist durch die mathematischen Zusammenhänge bestimmt, also zwangsläufig. Der Spin der Quarks muß $J = 1/2$ sein, denn nur so lassen sich sowohl Baryonen mit halbzahligem Spin als auch Mesonen mit ganzzahligem Spin aus Quarks aufbauen. Die präzisierte Vorstellung ist, daß Mesonen aus einem Quark und einem Antiquark, Baryonen aus drei Quarks bestehen. Die Baryonzahl der Quarks muß folglich $B = 1/3$ sein. Demnach besteht z. B. das Pion π^+ aus dem Quark u und dem Antiquark $\bar{d}$, kurz ausgedrückt $\pi^+ = u\bar{d}$, andere Beispiele sind $K^+ = u\bar{s}$, $p = uud$, $n = udd$. Außer der mathematischen Richtigkeit dieser Konstruktion gibt es auch experimentelle Hinweise, die zugunsten der Quarkvorstellung sprechen (↑ 3.2.5.3.).

Über die Massen der Quarks gehen die Ansichten auseinander, denn die Quarks sind bislang nicht als freie Teilchen beobachtet worden. Eine ziemlich triviale Erklärung dafür wäre, daß sie sehr schwer sind, bei der Bindung in Hadronen also ein starker Massendefekt auftritt, so daß die mit den verfügbaren Beschleunigern erreichbaren Energien nicht ausreichen, um sie als einzelne Teilchen freizusetzen (s. a. 3.2.6.2.). In jedem Fall sollte sich die Brechung der unitären Symmetrie auch in beträchtlichen Massendifferenzen innerhalb des Tripletts auswirken.

Allerdings ist das Quarkmodell viel komplizierter, als es dieser grobe Abriß zeigen könnte. Gewichtige theoretische Argumente, z. B. die Antisymmetrie der quantenmechanischen Wellenfunktion der Baryonen, fordern, daß die Quarks durch eine zusätzliche Quantenzahl gekennzeichnet werden, die den Namen *Colour* erhalten hat. Jeder der Quarks, also etwa u, kann danach in drei verschiedenen Colourzuständen existieren, die ein Colourtriplett bilden. Im Unterschied dazu sollten die uns bekannten Hadronen Coloursinguletts sein, also nicht in verschiedenen Colourzuständen auftreten.

3.2.3.3. Neue Teilchen – mehr Quarks. Die Mannigfaltigkeit der neuen Teilchen mit ψ, ψ' (↑ Tab. 3.2.2-2), den Mesonen D^+, D^0 und F^+ und den Baryonen Λ_c^+ und Σ_c^{++} findet eine einleuchtende Erklärung, wenn ein viertes Quark c (↑ Tab. 3.2.3-3) eingeführt wird. Die Teilchen ψ und ψ', sowie auch die dazwischen liegenden Resonanzen $\varkappa$ lassen sich als Zustände des gebundenen Systems aus c und $\bar{c}$, des sogenannten Charmoniumsystems verstehen. Das gilt auch für die beobachteten charakteristischen Zerfälle, die unter Emission eines Photons verlaufen. Die neuen Mesonen und Baryonen deutet man als gebundene Zustände, die ein c-Quark enthalten, nämlich $D^+ = c\bar{d}$, $D^0 = c\bar{u}$, $F^+ = c\bar{s}$, $\Lambda_c^+ = udc$, $\Sigma_c^{++} = uuc$.

Die Masse von D^+ ist mit 1,868 GeV/c^2 zu groß, als daß ψ oder ψ' in D^+ und D^- zerfallen dürften. Die in ψ oder ψ' enthaltenen Quarks c und $\bar{c}$ sind also nicht mehr in den Zerfallsprodukten enthalten, sondern haben sich durch Paarvernichtung in andere Quarkpaare umgewandelt. Nach einer gut bestätigten heuristischen Regel sind die Wirkungsquerschnitte solcher Reaktionen vergleichsweise klein, woraus sich die geringen Resonanzbreiten von ψ und ψ' erklären. Argumente für das vierte Quarkteilchen c sind übrigens schon vor der Entdeckung der neuen Teilchen ins Feld geführt worden, weil anders die Existenz der neutralen Ströme (↑ 3.2.4.5.) schwer zu verstehen ist.

Sobald man mit vier Quarks operiert, ist es nur folgerichtig, die Symmetriegruppe von SU(3) auf SU(4) zu vergrößern (↑ Tab. 3.2.3-1). Die wichtigste Konsequenz ist, daß sich das Mesonoktett zu einem 15-Multiplett erweitert, und das Baryonoktett zu einem 20-Multiplett. Die Mesonen D^+, D^0 und F^+ sind offensichtlich gemeinsam mit Pionen und Kaonen in das 15-Multiplett einzuordnen, die Baryonen Λ_c^+ und Σ_c^{++} in das 20-Multiplett; die großen Massendifferenzen weisen jedoch auf eine starke Brechung der SU(4)-Symmetrie hin. Die

Tab. 3.2.3-3 Quantenzahlen der Quarks

Quark	B	Q	I	I_3	Y	C	S
u	1/3	2/3	1/2	+1/2	1/3	0	0
d	1/3	−1/3	1/2	−1/2	1/3	0	0
s	1/3	−1/3	0	0	−2/3	0	−1
c	1/3	2/3	0	0	−2/3	1	0

neuen Resonanzen Y und Y' deuten auf einen weiteren Quark hin, der das Symbol b erhalten hat. Quantenzahlen können ihm aber erst dann zugeordnet werden, wenn andere Hadronen gefunden worden sind, die sich als gebundene Zustände unter Beteiligung von b oder $\bar{b}$ deuten lassen.

Eine Vergrößerung der Anzahl der fundamentalen Bausteine läuft eigentlich dem atomistischen Prinzip zuwider. Die künftige Erfahrung wird lehren, in welcher Form es in der Elementarteilchenphysik anwendbar ist.

3.2.4. Hierarchie der Wechselwirkungen

Die Wechselwirkungen zwischen den Elementarteilchen lassen sich zuerst einmal nach ihrer Stärke ordnen, und zum anderen nach dem Ausmaß, in dem die Erhaltungssätze respektiert werden. Beide Gesichtspunkte führen auf ein und dasselbe Schema (Tab. 3.2.4-1).

Tab. 3.2.4-1 Überblick über die Wechselwirkungen zwischen Elementarteilchen

	Stärke	Reichweite in cm	gute Quantenzahlen
stark	10	10^{-13}	$B, L_e, L_\mu, Q, Y, I_3, C, P, I$
elektromagnetisch	10^{-2}	∞	$B, L_e, L_\mu, Q, Y, I_3, C, P$
schwach	10^{-14}	$< 10^{-15}$	B, L_e, L_μ, Q

3.2.4.1. Erhaltungssätze. Eine physikalische Größe genügt einem Erhaltungssatz, wenn bei jeder Reaktion ihre Werte in Anfangs- und Endzustand übereinstimmen.

Der Erhaltungssatz für *Energie* und *Impuls* wird als absolut gültig angesehen, da noch in keinem Falle eine Verletzung beobachtet worden ist. Energie und Impuls eines Systems aus mehreren Teilchen setzen sich additiv aus den Größen der einzelnen Teilchen zusammen. Nach dem relativistischen Ausdruck für die Energie eines Teilchens mit Masse M und Impuls $\boldsymbol{p}$ (↑ 2.1.3.5.) besteht zwischen Energie und Impuls die Beziehung $E^2 - p^2c^2 = M^2c^4$. Energie- und Impulssatz werden benutzt, um Resonanzen anhand ihrer Zerfallsprodukte zu identifizieren. Als Beispiel soll der Zerfall $\varrho^0 \to \pi^+ + \pi^-$ dienen. Die Impulse $\boldsymbol{p}_1$ und $\boldsymbol{p}_2$ der Pionen werden gemessen; daraus können dann die Energien E_1 und E_2 berechnet werden, sobald die Teilchen als Pionen identifiziert worden sind (↑ 3.2.1.2.1.). Der Ausgangszustand hat die Energie $E = E_1 + E_2$ und den Impuls $\boldsymbol{p} = \boldsymbol{p}_1 + \boldsymbol{p}_2$, woraus sich die Masse zu $M^2 = (E^2 - p^2c^2)/c^4$ ergibt. Das Kennzeichen einer Resonanz ist, daß sich die so berechneten Massen mit nicht allzu großer Streuung um einen mittleren Wert, in unserem Beispiel um die Masse M_ϱ, gruppieren.

Der Erhaltungssatz für den *Drehimpuls* sagt aus, daß im Ruhsystem der Resonanz der Drehimpuls der Zerfallsprodukte (im Falle der spinlosen Pionen ihr resultierender Bahndrehimpuls) gleich dem Spin der Resonanz ist. Erschwert wird die Anwendung dadurch, daß der Bahndrehimpuls nicht unmittelbar in jedem einzelnen Prozeß meßbar ist. Nach der Quantentheorie läßt er sich aber aus der Winkelverteilung der Reaktionsprodukte bestimmen. Auf diese Weise wird der Spin von Resonanzen und solchen Teilchen ermittelt, die nicht im freien Flug beobachtet werden können.

Die Erhaltung der *Parität* (↑ 3.2.1.6.) bedeutet in anschaulicher und experimentell direkt prüfbarer Fassung: Jeder Prozeß, der einer Spiegelungstransformation unterworfen wird, d. h. in einem Spiegel betrachtet wird, läuft mit denselben Wirkungsquerschnitten ab, wie der ursprüngliche Prozeß. Bei schwachen Wechselwirkungen gilt die Paritätserhaltung nicht, d. h., diese Prozesse sind nicht spiegelungsinvariant (s. a. 3.2.4.5.).

Die Erhaltungssätze der Baryonzahl B (↑ 3.2.1.4.1.), der Leptonzahlen L_e und L_μ (↑ 3.2.1.4.2.) und der elektrischen Ladung Q sind uneingeschränkt gültig. Hyperladung Y (↑ 3.2.1.4.3.) und Charm C (↑ 3.2.1.4.4.) bleiben bei starken und elektromagnetischen Wechselwirkungen erhalten, nicht aber bei schwachen. Die Werte B, L_e, L_μ, Q, Y und C, die einem Mehrteilchensystem zugeordnet sind, setzen sich einfach additiv aus den Werten der einzelnen Teilchen zusammen. Als Beispiel soll das Baryon Ω^- mit $Y = -2$ dienen. Ein Zerfall unter Einfluß der starken Wechselwirkung könnte wegen des Erhaltungssatzes für Y nur in $\Xi^0 + K^-$ oder $\Sigma^+ + K^- + K^-$ erfolgen. Beides ist wegen des Erhaltungssatzes für die Energie nicht möglich. Folglich zerfällt Ω^- unter Einfluß der schwachen Wechselwirkung und ist daher langlebig.

Der *Isospin* bleibt bei starken Wechselwirkungen erhalten. Die Addition der Isospins der einzelnen Teilchen hat allerdings analog zu den quantentheoretischen Regeln für die Addition von Drehimpulsen (↑ 2.5.3.) zu erfolgen. So ist z. B. der Isospin des aus Pion ($I = 1$) und Nukleon ($I = 1/2$) bestehenden Systems entweder $1 + 1/2 = 3/2$ oder $1 - 1/2 = 1/2$, d. h., man findet Resonanzen mit $I = 3/2$ oder $I = 1/2$. Aus der Erhaltung des Isospins lassen sich auch Aussagen über den Zerfall von Resonanzen machen, z. B. daß die Zerfälle $\Delta^0 \to \pi^- + p$ und $\Delta^0 \to \pi^0 + n$ mit der relativen Häufigkeit 1 : 2 ablaufen.

Quantenzahlen, für die ein Erhaltungssatz gilt, heißen *gute Quantenzahlen.* Ganz allgemein entspricht jedem Erhaltungssatz eine Symmetrie oder Transformationsinvarianz. Unter einer Isospintransformation gehen die Mitglieder eines Isospinmultipletts, etwa des Dubletts (p, n) ineinander über. Eine Invarianz gegenüber dieser Transformation hätte die Konsequenz, daß die Massen innerhalb eines Multipletts übereinstimmen. Die Größe der Massendifferenzen (↑ Tab. 3.2.2-1) ist ein Maß für die Stärke der Symmetriebrechung.

3.2.4.2. Feynmangraphen. Sie haben ihren Ursprung im Formalismus der relativistischen Quantenfeldtheorie und sind eigentlich symbolische Vorschriften zur Berechnung von Wirkungsquerschnitten. Die Feynmangraphen haben zugleich aber auch einen sehr anschaulichen Gehalt, indem sie unsere Vorstellungen von den Elementarteilchen und ihren Wechselwirkungsprozessen in einer bildhaften Form darstellen. Ein Graph besteht aus *Linien* und *Vertices*, d. h. Knotenpunkten. Den Linien sind Teilchen zugeordnet, insbesondere bedeuten Linien mit Pfeil Teilchen mit dem Spin $\frac{1}{2}$, eine Wellenlinie das Photon, und durchbrochene Linien Mesonen. Den Linien ist auch ein Impuls zugeordnet, der in einen Vertex entweder einläuft oder ausläuft.

Wenn Pfeilrichtung der Linie und Richtung des zugehörigen Impulses übereinstimmen, beschreibt die Linie ein Teilchen, sonst das Antiteilchen. Die Vertices, in denen sich im allgemeinen drei oder vier Linien treffen, symbolisieren Wechselwirkungsakte (Abb. 3.2.4-2a). An jedem Vertex sind die Erhaltungssätze erfüllt, die für die betreffende Wechselwirkung gelten, d. h., Energie, Impuls und Ladung bleiben in jedem Fall erhalten. Die zwei folgenden Punkte verdienen besondere Beachtung.

Ein Graph beschreibt immer eine Gruppe von Prozessen, die durch die Ersetzung „einlaufendes Teilchen → auslaufendes Antiteilchen" und umgekehrt auseinander hervorgehen. So beschreibt z. B. Abb. 3.2.4-2b den Prozeß $p \to n + \pi^+$, wenn die Impulse von Neutron und Pion aus dem Vertex laufen, dagegen den Prozeß $p + \bar{n} \to \pi^+$, wenn die Impulse von Proton und Antineutron in den Vertex einlaufen.

Die inneren Linien eines Graphen, die also zwei Vertices verbinden, stellen *virtuelle Teilchen* dar. Sie sind im Unterschied zu den durch äußere Linien repräsentierten Teilchen nicht beobachtbar, und es gilt auch nicht die übliche Relation zwischen Energie, Impuls und Masse.

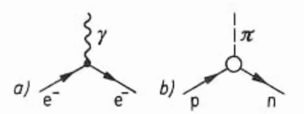


Abb. 3.2.4-2 *a)* fundamentaler Elektron-Photon-Vertex, *b)* Pion-Nukleon-Vertex (Impulse sind nicht eingezeichnet)

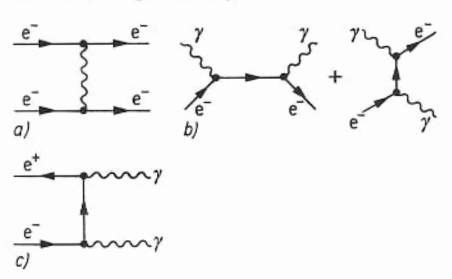


Abb. 3.2.4-3 Feynmangraphen; *a)* zur Streuung von Elektronen aneinander, *b)* zur Comptonstreuung an Elektronen, *c)* zur Zerstrahlung $e^- + e^+ \to \gamma + \gamma$

Die Wechselwirkungen und Kräfte zwischen reellen Teilchen beruhen auf dem Austausch von virtuellen Teilchen. Die virtuellen Teilchen spielen schon in den einfachsten Prozessen der Quantenelektrodynamik, bei der Elektron-Elektron-Streuung (Abb. 3.2.4-3a), der Comptonstreuung (Abb. 3.2.4-3b) und der Elektron-Positron-Zerstrahlung (Abb. 3.2.4-3c) eine Rolle. Der Austausch eines virtuellen Photons zwischen e^- und e^-, e^- und e^+, e^+ und e^+ erzeugt die Coulombkraft. Die Kernkräfte (↑ 3.3.4.), d. h. die zwischen den Nukleonen wirkenden Kräfte, haben ihren Ursprung im Austausch von virtuellen Pionen, ρ-Mesonen usw. (Abb. 3.2.4-4a). Da es sich hier um den Austausch von massiven virtuellen Teilchen

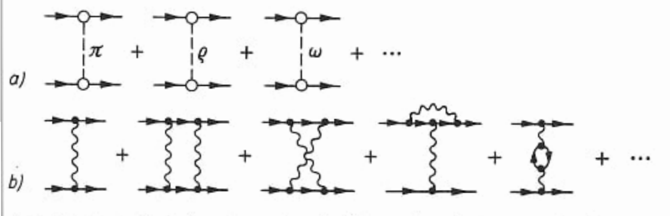


Abb. 3.2.4-4 Beiträge a) zu den Kräften, die zwischen den Nukleonen wirken, b) zu den elektromagnetischen Kräften zwischen Elektronen

handelt, ergeben sich die kurzreichweitigen Kernkräfte.

Jeder Vertex trägt eine Kopplungskonstante, die ein Maß für die Stärke der Wechselwirkung ist. Der *Elektron-Photon-Vertex* ist proportional zur Elementarladung e. Folglich ist die Kraft zwischen zwei Elektronen bzw. Elektron und Positron proportional zu $(-e)\,(-e) = e^2$ bzw. $(-e)\,(+e) = -e^2$, d. h. einmal abstoßend und das andere Mal anziehend. Als Maß für die Stärke der elektromagnetischen Wechselwirkung benutzt man den dimensionslosen Ausdruck $e^2/4\pi\varepsilon_0\hbar c \approx 1/137$, die sogenannte Sommerfeldsche Feinstrukturkonstante (s. a. 4.2.2.6.). Der *Pion-Nukleon-Vertex* und allgemein die Meson-Baryon-Vertices sind nur indirekt und nicht sehr genau bestimmbar. Die zur Feinstrukturkonstante analogen Ausdrücke liegen in der Größenordnung von 10.

Grundsätzlich tragen zu jedem Prozeß Graphen beliebig hoher Ordnung bei (Abb. 3.2.4-4b). Die sukzessive Berücksichtigung von Graphen immer höherer Ordnung heißt Störungsrechnung. Wegen der Kleinheit der Elektron-Photon-Kopplung stellt der erste Graph den weitaus größten Beitrag dar. Das ist der Grund dafür, daß mit großer Genauigkeit die Kraft zwischen Elektronen, und ähnlich die Kraft zwischen Elektron und Proton, durch die einfache Coulombkraft ersetzt werden kann. Diesem glücklichen Umstand ist es zu verdanken, daß die Quantentheorie weitgehende Einsichten in die Gesetzmäßigkeiten der Atomhülle gewinnen konnte. Die höheren Graphen sind in der Quantenelektrodynamik berechnet worden. Sie führen auf die Strahlungskorrekturen der Kräfte, die sich in kleinen Verschiebungen der Energieniveaus der Atomhülle auswirken (↑ 4.2.2.6.). Auch zum magnetischen Moment μ (↑ 3.2.1.3.5.) des Elektrons tragen Graphen beliebig hoher Ordnung bei (Abb. 3.2.4-5). Der erste Graph, also der fundamentale Vertex, liefert das normale magnetische Moment, die übrigen Graphen den anormalen Beitrag. Die berechneten Strahlungskorrekturen stimmen ausgezeichnet mit den experimentellen Fakten überein.



Abb. 3.2.4-5 Beiträge zum Elektron-Photon-Vertex

Im Unterschied dazu ist die starke Wechselwirkung zwischen Nukleonen wegen der Größe der Kopplungskonstante nicht auf dem Wege der Störungsrechnung zu behandeln. Es trägt in jedem Fall eine quantitativ nicht überschaubare Reihe von Graphen bei (↑ Abb. 3.2.4-4a). Daher konnten die Kernkräfte bisher nicht in ähnlicher Weise berechnet werden wie die elektromagnetischen Kräfte.

Der Methode der Feynmangraphen haftet eine prinzipielle Schwierigkeit an. Gewisse Graphen mit geschlossenen Linienzügen (Abb. 3.2.4-6) ergeben bei der Berechnung ein unendliches Resultat. Zur Lösung dieses Problems ist die mathematisch sehr anspruchsvolle Methode der *Renormierung* entwickelt worden. Bei diesem Verfahren werden die experimentell gemessenen Werte von *Kopplungskonstante* und *Elektronenmasse* in der Theorie als Differenzen divergierender Größen dargestellt. Berechenbar sind nur Strahlungskorrekturen und Wirkungsquerschnitte.

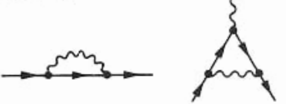


Abb. 3.2.4-6 Beispiele für divergierende Graphen

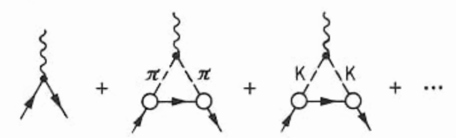


Abb. 3.2.4-7 Beiträge zum Proton-Photon-Vertex; im Falle des Neutrons fehlt der erste Graph

3.2.4.3. Starke Wechselwirkung. Die starke Wechselwirkung beherrscht die Streu- und Erzeugungsprozesse von Hadronen. Auch die großen anormalen Beiträge zum magnetischen Moment der Baryonen sind ein Effekt der starken Wechselwirkung (Abb. 3.2.4-7). Man spricht in diesem Zusammenhang von einer Wolke virtueller Mesonen rund um das Baryon. Die Stärke der Wechselwirkung drückt sich in Meson-Baryon-Kopplungen der Größenordnung 10 aus. Man wird diese Kopplungen nicht in demselben Maße als fundamental ansehen können, wie etwa die Elektron-Photon-Kopplung. Die starke Wechselwirkung respektiert alle Erhaltungssätze und Symmetrien mit Ausnahme der unitären. Die Brechung der unitären Symmetrien äußert sich in den relativ großen Massendifferenzen eines Supermultipletts.

3.2.4.4. Elektromagnetische Wechselwirkung. Sie kann im Falle der Leptonen auf den fundamentalen Elektron-Photon-Vertex (↑ Abb. 3.2.4-5) und einen analogen für Myonen zurückgeführt werden. Anders bei den Hadronen, wo die elektromagnetische stets mit der

starken Wechselwirkung zusammenwirkt. Ein einfaches Beispiel bietet der Proton-Photon-Vertex (↑ Abb. 3.2.4-7). Zerfallsprozesse, die unter Beteiligung der elektromagnetischen Wechselwirkung ablaufen, zeichnen sich durch Lebensdauern von der Größenordnung 10^{-16} s aus. Beispiele sind $\Sigma^0 \rightarrow \Lambda + \gamma$ und $\pi^0 \rightarrow 2\gamma$. Wenn dem Photon die Hyperladung $Y = 0$ zugeordnet wird, kann man von einem Erhaltungssatz für Y reden. Beide Prozesse sind auch dadurch bemerkenswert, daß Photonen mit elektrisch neutralen Teilchen wechselwirken. Auf der Grundlage des Quarkmodells kann man sich das so erklären, daß es eigentlich die virtuellen Quarks sind, an die die Photonen koppeln (Abb. 3.2.4-8).

Ein Prozeß, bei dem Y erhalten bleibt, I jedoch nicht, wird im erweiterten Sinne eine elektromagnetische Wechselwirkung genannt, auch wenn kein reelles Photon daran beteiligt ist. Ein Beispiel bietet der Zerfall des η-Mesons in drei Pionen. Ein elektromagnetischer Effekt ist auch die Massenaufspaltung innerhalb der Isospinmultipletts. Sie bleibt im Baryonoktett geringer als 0,01 GeV/c^2, das ist eine Größenordnung kleiner als der Abstand der Multipletts untereinander (↑ Tab. 3.2.2-1).

3.2.4.5. Schwache Wechselwirkung. Ein relativ einfaches Bild bietet die schwache Wechselwirkung der Leptonen untereinander. Es gibt einen fundamentalen Vertex, der über eine *Vier-Fermionen-Kopplung* entweder den Myonzerfall $\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$ (Abb. 3.2.4-9a) oder etwa den Streuprozeß $\bar{\nu}_\mu + e^+ \rightarrow \mu^+ + \bar{\nu}_e$ (Abb. 3.2.4-9b) beschreibt. Die Kopplungskonstante, die mit dem Pion-Nukleon-Vertex zu vergleichen wäre, ist im Falle der schwachen Wechselwirkung klein; $G \approx 1{,}1 \cdot 10^{-5}/M_p^2$. Die schwache Wechselwirkung zwischen Hadronen führt zu weit komplizierteren Erscheinungen, insbesondere bei den Zerfällen. Charakteristisch ist die große Zahl von *Zerfallskanälen* (Abb. 3.2.4-10). Die Verzweigungsverhältnisse geben die Wahrscheinlichkeit an, daß der Zerfall in dem betreffenden Kanal stattfindet. Man hat zwischen leptonischen und nichtleptonischen Zerfällen zu unterscheiden. Leptonische Zerfälle gehen unter Aussendung von Leptonen einher, z. B. der β-Zerfall $n \rightarrow p + e^- + \bar{\nu}_e$ und der Zerfall des Pions $\pi^+ \rightarrow \mu^+ + \nu_\mu$. Beispiele für nichtleptonische Zerfälle sind $\Lambda \rightarrow p + \pi^-$ oder $\Sigma^- \rightarrow n + \pi^-$. Die Größe der Lebensdauer von etwa 10^{-10} s zeugt davon, daß auch bei diesen Prozessen die schwache Wechselwirkung beteiligt ist, obwohl keine Neutrinos ausgesendet werden. In der Regel ändert sich bei nichtleptonischen Zerfällen die Hyperladung um $\Delta Y = 1$, und der Isospin um $\Delta I = \frac{1}{2}$. Die Graphen dieser Zerfallsprozesse wird man sicher nicht als fundamentale Vertices ansehen können. Fundamentale Kopplungen setzt man statt dessen zwischen Quarks und Leptonen an (Abb. 3.2.4-11). Auf diese Kopplungen kann auch der Prozeß des β-Zerfalls zurückgeführt werden (Abb. 3.2.4-12). Nichtleptonische Prozesse enthalten die Leptonen nur als virtuelle Teilchen.

Die Besonderheit der schwachen Wechselwirkungen besteht darin, daß die *Parität nicht*

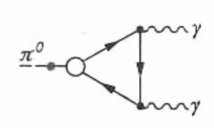


Abb. 3.2.4-8 Mechanismus des neutralen Pionzerfalls; die Photonen koppeln an virtuelle Quarks

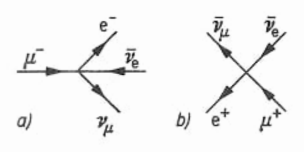


Abb. 3.2.4-9 *a)* Fundamentale Kopplung des Myonzerfalls, *b)* Graph für den Prozeß $\bar{\nu}_\mu + e^+ \rightarrow \mu^+ + \bar{\nu}_e$

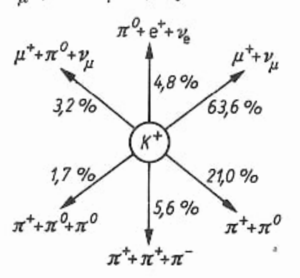


Abb. 3.2.4-10 Zerfallsmöglichkeiten eines positiven K-Mesons mit Angabe der Zerfallswahrscheinlichkeit in die einzelnen Kanäle

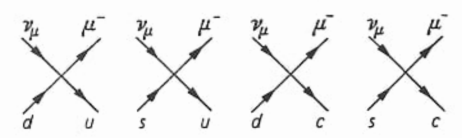


Abb. 3.2.4-11 Fundamentale Kopplungen der schwachen Wechselwirkung im Quarkmodell; entsprechend für e^- und ν_e

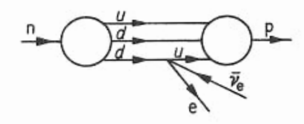


Abb. 3.2.4-12 Zurückführung des β-Zerfalls auf die fundamentale Kopplung zwischen Quarks und Leptonen

erhalten bleibt (s. a. 3.2.4.1.). Soweit es sich um leptonische Prozesse handelt, ist der Grund leicht einzusehen: Bei der Spiegelungstransformation geht das linkshändige Neutrino in ein rechtshändiges über, das es als Neutrino nicht gibt, wodurch ein physikalisch möglicher Prozeß in einen unmöglichen transformiert wird. Eine Transformation, welche Neutrinos in real existierende Teilchen, d. h. Antineutrinos überführt, ist durch das Produkt aus Spiegelung P und Ladungskonjugation C (C führt Teilchen in Antiteilchen über) gegeben. Unter dieser kombinierten Operation $C\,P$ sollte die schwache Wechselwirkung invariant sein. Das ist auch der Fall, bis auf eine einzige Ausnahme.

Beim Zerfall der *neutralen Kaonen* (↑ 3.2.2.2.) wird die $C\,P$-Invarianz verletzt. Wäre das nicht so, dann dürfte der langlebige Zustand K_L^0 wohl in drei Pionen übergehen, nicht aber in zwei. Das experimentelle Ergebnis ist, daß 0,2% aller Zustände K_L^0 in ein Paar geladener Pionen zerfällt. Auch die äußerst kleine Differenz zwischen den Massen von K_L^0 und K_S^0 in Höhe von etwa $3 \cdot 10^{-12}$ MeV/c^2 geht auf die $C\,P$-verletzende Wechselwirkung zurück, woraus folgt, daß diese offenbar sehr schwach ist.

In den letzten Jahren hat noch eine andere Entdeckung von sich reden gemacht, die unter dem Schlagwort „*Neutrale Ströme*" bekannt geworden ist. Es handelt sich hierbei um eine ganz neue Klasse von Prozessen, die die Einführung von zusätzlichen fundamentalen Vertices erforderlich macht. Es wurden Reaktionen beobachtet, die durch Neutrinos ν_μ oder Antineutrinos $\bar{\nu}_\mu$ ausgelöst werden, ohne daß im Endzustand ein Myon oder Antimyon zu finden war. Die normalen Reaktionen lauten dagegen $\nu_\mu + \mathrm{n} \to \mathrm{p} + \mu^-$ bzw. $\bar{\nu}_\mu + \mathrm{p} \to \mathrm{n} + \mu^+$. Diese Experimente bewiesen, daß es auch die Reaktionen $\nu_\mu + \mathrm{n} \to \nu_\mu + \mathrm{n}$ und $\bar{\nu}_\mu + \mathrm{p} \to \bar{\nu}_\mu + \mathrm{p}$ gibt. Der Umstand, daß das Lepton im Gegensatz zu den normalen Reaktionen seine elektrische Ladung nicht ändert, führte zu der Bezeichnung neutrale Ströme. Besonders eindrucksvoll wurde die Existenz zusätzlicher Kopplungen durch die Beobachtung des Prozesses $\bar{\nu}_\mu + \mathrm{e}^- \to \bar{\nu}_\mu + \mathrm{e}^-$ belegt. Nach landläufigen Vorstellungen wäre dieser Prozeß auszuschließen, da Antineutrinos $\bar{\nu}_\mu$ nur mit Positronen reagieren (↑ Abb. 3.2.4-9b). Die Entdeckung der neuen Prozesse ist darum so bedeutsam, weil ihre Existenz eine unerläßliche Voraussetzung für das Programm einer einheitlichen theoretischen Behandlung der elektromagnetischen und schwachen Wechselwirkungen ist (s. a. 3.2.6.1.).

3.2.4.6. Gravitative Wechselwirkung. Ohne Zweifel unterliegen auch die Elementarteilchen der Gravitation, d. h. der universellen gegenseitigen Anziehung aller Massen. Der Effekt ist allerdings äußerst klein, so daß man bei gegenwärtigen Experimenten im Laboratorium völlig von der Gravitation absehen kann. Die Schwäche dieser Wechselwirkung wird ersichtlich, wenn man die zwischen zwei Protonen wirkende elektrische Coulombkraft mit der Massenanziehung vergleicht. Es ergibt sich das Verhältnis $(e^2/4\pi\varepsilon_0)/\varkappa M_\mathrm{p}^2 \sim 10^{36}$. Für die Gravitationskonstante $\varkappa$ gilt $\varkappa \approx 6{,}67 \cdot 10^{-11}\,\mathrm{m^3kg^{-1}s^{-2}}$. Bei sehr hohen Energien wird jedoch die Gravitation nicht immer zu vernachlässigen sein. Das liegt anschaulich ausgedrückt daran, daß auch die Energie entsprechend der Einsteinschen Beziehung $m = E/c^2$ gravitierend wirkt, und somit eine hinreichende Vergrößerung der Teilchenenergie schließlich die Kleinheit der Gravitationskonstante zu kompensieren vermag. Auch Probleme der kosmischen Physik, z. B. die Schwarzen Löcher und ihr Zerfall, weisen auf eine enge Verflechtung von Elementarteilchen und Gravitation hin. Eine konsistente theoretische Behandlung beider Gebiete, die eine Quantentheorie der Gravitation einschließen muß, ist eine Aufgabe von grundlegender Bedeutung.

3.2.5. Struktur der Hadronen

Aus der Vielzahl von Reaktionen, die bisher experimentell untersucht worden sind, ragen drei Gruppen heraus, die eine besondere Bedeutung für die Strukturbestimmung der Hadronen erlangt haben. Diesen Experimenten ist gemeinsam, daß sie hochenergetische Leptonen verwenden.

3.2.5.1. Elastische Elektronenstreuung. Die räumliche Verteilung von elektrischer Ladung und magnetischem Moment von Nukleonen wird mit Hilfe der elastischen Streuung von Elektronen (Abb. 3.2.5-1) gemessen. Das erfolgt über die *elektromagnetischen Formfaktoren*, die der Abweichung von einem punktförmigen Teilchen Rechnung tragen, und die unmittelbar aus den Wirkungsquerschnitten

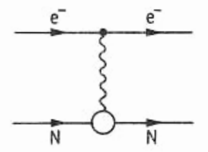


Abb. 3.2.5-1 Elastische Elektron-Nukleon-Streuung

zu bestimmen sind. Der dynamische Ursprung der Formfaktoren wird durch den Feynmangraph in Abb. 3.2.4-7 beschrieben. Es ist bei diesen Experimenten von wesentlicher Bedeutung, daß sich die Sonde, das Elektron, hinsichtlich seiner Eigenschaften deutlich vom Objekt, dem Nukleon, unterscheidet. Das Elektron kann nämlich nur durch den Aus-

tausch virtueller Photonen mit dem Nukleon wechselwirken. Wegen der Kleinheit der elektromagnetischen Kopplungskonstante darf man sicher sein, daß der Prozeß hervorragend durch den Austausch eines einzigen Photons beschrieben wird. Das ist letztlich der Grund dafür, daß nur die Formfaktoren der Nukleonen, und im Prinzip auch die der anderen Hadronen experimentell zu bestimmen sind, nicht aber die Formfaktoren der Leptonen.

Die Formfaktoren, und damit die räumliche Ausdehnung (↑ Abb. 3.2.1-3), werden durch elastische Streuprozesse gemessen, bei welchen das Nukleon ein Nukleon bleibt. Wie bei einer Vermessung makroskopischer Körper bleibt die Struktur des Objekts unangetastet. Dagegen erfolgt die Untersuchung der Bestandteile der Hadronen auf dem Wege der tief-inelastischen Streuung (↑ 3.2.5.3.) derart, daß das Objekt zerstört wird.

3.2.5.2. Elektron-Positron-Vernichtung. Dieser Prozeß ist von besonderer Bedeutung, da die

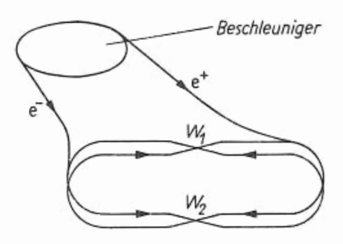


Abb. 3.2.5-2 Schema von Speicherringen; aus dem Beschleuniger werden Elektronen und Positronen in die vertikal versetzten Speicherringe eingespeist; die Strahlen treffen in W_1 und W_2 aufeinander

Mehrzahl der neuen Teilchen (↑ 3.2.2.2.) auf diesem Wege entdeckt wurde, und auch in Zukunft überraschende Entwicklungen nicht auszuschließen sind. Außerdem eröffnet die Elektron-Positron-Vernichtung die Möglichkeit, Aufschluß über die Zahl der Quarks zu erhalten. Es handelt sich somit ebenfalls um ein Experiment, das mit der Aufklärung der Hadronstruktur zu tun hat. Der Prozeß der Paarvernichtung wird an den Elektron-Positron-Speicherringen untersucht (Abb. 3.2.5-2) (↑ 3.1.2.5.1., 3.1.2.5.2.).

Der dynamische Mechanismus der Paarvernichtung ist dadurch gekennzeichnet, daß das Leptonpaar nur durch virtuelle Photonen mit den Teilchen des Endzustands in Wechselwirkung stehen kann. Wegen der Kleinheit der elektromagnetischen Kopplungskonstanten ist es der Austausch eines einzigen Photons, der den Prozeß in guter Näherung beschreibt. Die Erzeugung von Hadronen stellt man sich im Quarkmodell (↑ 3.2.3.2.) so vor, daß das virtuelle Photon an ein virtuelles Quark-Antiquark-Paar koppelt, das schließlich in einen reellen Hadronzustand übergeht (Abb. 3.2.5-3).

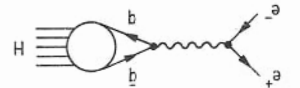


Abb. 3.2.5-3 Erzeugung von Hadronen H durch Elektron-Positron-Vernichtung

Diese Vorstellung führt zu einer quantitativen theoretischen Aussage über den Quotienten aus dem Wirkungsquerschnitt der Hadronerzeugung und dem Wirkungsquerschnitt der Reaktion $e^- + e^+ \rightarrow \mu^- + \mu^+$. Dieser Quotient R sollte für große Energien gegen den Wert $\sum Q_i^2$ streben. Hierbei bedeuten Q_i die elektrischen Ladungen der Quarks, gemessen in Einheiten der Elementarladung. Bei Berücksichtigung der dort aufgeführten Quarks findet man $\sum Q_i^2 = 3 \cdot \{4/9 + 1/9 + 1/9 + 4/9\} = 10/3$. Der Faktor 3 trägt dem Umstand Rechnung, daß jeder Quark in den drei verschiedenen Colourzuständen auftreten kann. Dagegen wächst bei einer Energie von etwa 28 GeV der gemessene Quotient R auf den Wert $R \approx 4$ an. Daraus hat man den Schluß zu ziehen, daß die Tabelle der Quarks noch nicht vollzählig ist.

3.2.5.3. Tief-inelastische Streuung von Leptonen. Diese Experimente dienen der eigentlichen Erforschung der Hadronstruktur. Praktisch handelt es sich um die Streuung von Myonen oder Myon-Neutrinos an Nukleonen, wobei sowohl die Energie E der einfallenden Leptonen (100 GeV und mehr), als auch die sogenannte *Impulsübertragung* auf das Nukleon groß sind. Hierbei ist die Impulsübertragung $\sqrt{Q^2}$ durch die relativistische Formel $Q^2 = (\boldsymbol{p} - \boldsymbol{p}')^2 - (E - E')^2/c^2$ bestimmt, wobei $\boldsymbol{p}'$ und E' Impuls bzw. Energie des gestreuten Leptons bezeichnen. Die Umschreibung „tief-inelastisch" drückt zuerst einmal aus, daß der Prozeß ein wesentlich inelastischer ist: das Nukleon geht im allgemeinen in ein komplexes System von Hadronen über. Ganz anders als bei der Bestimmung der räumlichen Ausdehnung (s. a. 3.2.5.1.) wird das Objekt der Strukturbestimmung zerstört. Die großen Werte der Impulsübertragung ermöglichen den Einblick in die

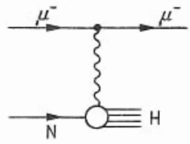


Abb. 3.2.5-4 Tief-inelastische Myon-Nukleon-Streuung

innere Struktur der Hadronen. Auch bei diesem Prozeß kann die elektromagnetische oder schwache Wechselwirkung der Leptonen näherungsweise auf die niedrigste Ordnung der Kopplungskonstanten beschränkt werden. Im Falle der Streuung von Myonen wird der Prozeß demnach durch den Feynmangraph (Abb. 3.2.5-4) beschrieben.

Die bisher durchgeführten Experimente zeigten ein überraschendes Ergebnis. Es wurden Wirkungsquerschnitte beobachtet, wie sie am ehesten bei der Streuung von Leptonen an punktförmigen Teilchen zu erwarten wären. Darüber hinaus ist ein Rückschluß auf den Spin dieser Streuzentren möglich; er beträgt $J = 1/2$. Die erste Schlußfolgerung ist, daß der tief-inelastische Streuprozeß nicht am Hadron als Ganzem erfolgt, sondern an irgendwelchen Bestandteilen des Hadrons, die man, vorerst ohne genaue Kenntnis ihrer Natur, als *Partonen* bezeichnet. Eine Streuung am Nukleon als Ganzem liegt bei der elastischen Streuung vor; dort zeigt sich sehr deutlich, daß das Nukleon nicht punktförmig ist (s. a. 3.2.5.1.). Als zweites liegt die Vermutung sehr nahe, die Partonen mit den Quarks zu identifizieren. Im Prinzip kann auch diese Frage durch das Experiment beantwortet werden. Die spezifischen fundamentalen Kopplungen zwischen Quarks und Photonen bzw. Quarks und Leptonen (↑ Abb. 3.2.4-11) führen nämlich zu bestimmten Beziehungen zwischen den Wirkungsquerschnitten der Streuung von Myonen, Neutrinos und Antineutrinos an Nukleonen, den sogenannten *Summenregeln.* Die bisher vorliegenden experimentellen Ergebnisse lassen innerhalb ihrer Fehlergrenzen den Schluß zu, daß die Partonen nichts anderes als Quarks sind. Es ist auch möglich, die Quarkanteile quantitativ durch Verteilungsfunktionen

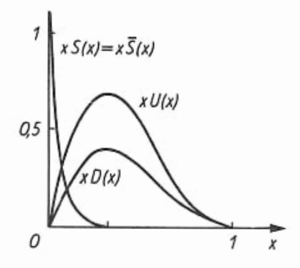


Abb. 3.2.5-5 Gegenwärtige Vorstellungen von den Quarkverteilungsfunktionen des Protons

zu beschreiben. Man stelle sich ein Nukleon mit dem Impuls p vor. Zu diesem Gesamtimpuls tragen die Quarks mit dem Bruchteil x, $(0 < x < 1)$, bei. Der Anteil der einzelnen Quarksorten u, d usw. kann dann durch die statistischen Verteilungsfunktionen $U(x)$, $D(x)$ usw. erfaßt werden. Zum Beispiel gibt $U(x)\,\Delta x$ die Wahrscheinlichkeit an, im Nukleon ein u-Quark mit einem zwischen xp und $(x + \Delta x)\,p$ liegenden Impuls zu finden. Abb. 3.2.5-5 zeigt Verteilungsfunktionen, die mit den experimentellen Resultaten verträglich sind. Ein eindeutiger Schluß auf die einzelnen Kurven ist allerdings nicht möglich. Es gibt Hinweise darauf, daß diese Vorstellungen noch zu naiv sind. Der deutlichste ist die Bilanz für den Nukleonimpuls, die unter Zuhilfenahme der gegebenen Quarkverteilungsfunktionen auf ein Defizit von etwa 50% führt! Eine befriedigende Erklärung für diese und andere Besonderheiten wird durch die *Quantenchromodynamik* (↑ 3.2.6.2.) erbracht, die eine spezielle Quantenfeldtheorie von „farbigen" Quarks (↑ 3.2.3.2.) und Gluonen (↑ 3.2.6.2.) ist.

War die Untersuchung der tief-inelastischen Streuung schon in den letzten Jahren von erstrangigem Einfluß auf die Herausbildung physikalischer Vorstellungen über die Hadronen, so wird ihre Bedeutung in den nächsten Jahren eher noch zunehmen. Das kommt auch schon in der Größe der experimentellen Anlagen zum Ausdruck. Das 1979 im Kernforschungszentrum in Genf angelaufene Experiment verwendet ein Myonspektrometer von mehr als 50 m Länge, das unter anderem aus einigen Hundert Zählerkammern und 1600 t Magneteisen besteht. Man wird damit in Bereiche von $E = 280$ GeV und $Q^2 = 500\ [\text{GeV}]^2/c^2$ vordringen. Die tief-inelastische Streuung ist von allen Experimenten am besten geeignet, die in der Quantenchromodynamik formulierten Vorstellungen von den Hadronen und ihren Wechselwirkungen zu prüfen.

3.2.6. Ansätze zu einer Theorie der Elementarteilchen

Unter dem Begriff Theorie soll hier eine systematische mathematische Formulierung physikalischer Zusammenhänge verstanden werden. Man muß zwischen allgemeiner theoretischer Beschreibung und einer konkreten Theorie, etwa von der Art der Quantenelektrodynamik, unterscheiden. Der erstgenannte Zugang scheint weitgehend abgeschlossen zu sein. Darunter fallen die relativistische Quantentheorie für nicht-wechselwirkende Teilchen, die theoretische Analyse von Wirkungsquerschnitten, die Untersuchung von Symmetrietransformationen und Erhaltungssätzen. Tieferliegende Ergebnisse sind das Theorem über den Zusammenhang zwischen Spin und Statistik (Teilchen mit halbzahligem Spin sind *Fermionen*; Teilchen mit ganzzahligem Spin sind *Bosonen* ↑ 2.5.5.2.) und das *TCP-Theorem.* Es besagt, daß alle Wechselwirkungen unter der Kombination der

Symmetrietransformationen Spiegelung, Ladungskonjugation (s. a. 3.2.4.5.) und Zeitumkehr invariant sind. Ein anderes allgemeines Resultat bezieht sich auf das Verhalten der totalen Wirkungsquerschnitte von Hadronen bei hohen Energien E, sie können höchstens wie $(\log E)^2$ anwachsen. Ganz anders steht es um die Berechnung von Wirkungsquerschnitten, Lebensdauern oder Massen. Das kann nur auf der Grundlage einer konkreten Theorie erfolgen. Gegenwärtig dominieren quantenfeldtheoretische Zugänge nach dem Vorbild der Quantenelektrodynamik.

3.2.6.1. Einheitliche Theorie der schwachen und elektromagnetischen Wechselwirkungen. Das wichtigste Ziel dieser Theorie ist es, auch die schwache Wechselwirkung wie die elektromagnetische renomierbar zu formulieren (s. a. 3.2.4.2.). Man erreicht es durch eine weitgehend ähnliche Behandlung beider Wechselwirkungen; insbesondere bedeutet das eine

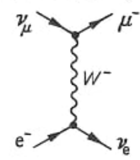


Abb. 3.2.6-1 Austausch eines intermediären Bosons anstelle der Vier-Fermionen-Kopplung

Abkehr von der fundamentalen Vier-Fermionen-Kopplung (↑ Abb. 3.2.4-9 und 3.2.4-11). Statt dessen denkt man sich die schwache Wechselwirkung ebenfalls durch den Austausch von virtuellen Teilchen mit dem Spin 1, den sogenannten *intermediären Bosonen* vermittelt (Abb. 3.2.6-1). Diese Teilchen müssen sehr schwer sein, denn anders wäre die neue Formulierung nicht mit der bei niedrigen und mittleren Energien gut bewährten Vier-Fermionen-Kopplung verträglich. Eine unerläßliche Bedingung für den Erfolg eines solchen Vorgehens sind die Kopplungen der neutralen Ströme (↑ 3.2.4.5.), die glücklicherweise im Experiment beobachtet worden sind. Die einfachste Variante ist die ursprüngliche, von SALAM und WEINBERG aufgestellte einheitliche Theorie. Die eigentliche schwache Wechselwirkung wird durch den Austausch von drei intermediären Bosonen W^+, W^- und Z^0 vermittelt, deren Massen größer als 80 GeV/c^2 sein sollten. Das altbekannte masselose Photon hat man ebenfalls als intermediäres Boson zu verstehen.

Obwohl es die derzeit verfügbaren Energien noch nicht gestatten, einen Effekt der schweren intermediären Bosonen wahrzunehmen oder sie zu erzeugen, kann schon jetzt die Konsistenz der einheitlichen Theorie in ganz unabhängigen Experimenten getestet werden. Die eine Gruppe von Experimenten zielt auf den Vergleich der Wirkungsquerschnitte von Prozessen ab, die entweder durch geladene oder durch neutrale Ströme vermittelt werden. Bei den anderen Experimenten geht es um die Bestimmung des relativen Anteils von Z^0-Austausch in Prozessen, die vorwiegend über Photonenaustausch ablaufen, z. B. bei der tief-inelastischen Streuung von Elektronen an Nukleonen (↑ Abb. 3.2.5-4). Da sich die Paritätsverletzung der schwachen Wechselwirkung in der neuen Formulierung durch paritätsverletzende Kopplungen zwischen intermediären Bosonen und Leptonen bzw. Quarks ausdrückt, handelt es sich um solche Fälle von Paritätsverletzung, die nach den älteren Vorstellungen nicht hätten auftreten dürfen. Dazu gehören auch die vieldiskutierten Effekte von *Paritätsverletzung in der Atomhülle.* Die verschiedenen Experimente müssen ein und denselben Wert des sogenannten Weinbergwinkels, des einzigen freien Parameters der Theorie, erbringen. Für eine direkte experimentelle Bestätigung der einheitlichen Theorie ist jedoch der Nachweis der intermediären Bosonen unerläßlich. Die in Genf im Bau befindlichen Proton-Antiproton-Speicherringe mit einer Energie von 2mal 270 GeV eröffnen die Möglichkeit der Erzeugung von intermediären Bosonen im Labor. Es erfordert allerdings einen experimentellen Aufwand von bisher nicht gekannter Größenordnung, diese Teilchen inmitten einer Vielzahl von gleichzeitig erzeugten Hadronen und Leptonen nachzuweisen.

3.2.6.2. Quantenchromodynamik. Schon in der Bezeichnung kommt zum Ausdruck, daß bei dieser Quantenfeldtheorie von Quarks (↑ 3.2.3.2.) die zusätzliche Quantenzahl Farbe oder Colour eine wesentliche Rolle spielt. Auch wird auf eine gewisse Verwandtschaft mit der Quantenelektrodynamik angespielt. Ähnlich wie der Austausch von Photonen die Wechselwirkung zwischen Elektronen vermittelt (s. a. 3.2.4.2.), führt man die Wechselwirkung zwischen Quarks auf den Austausch von *Gluonen* zurück, das sind masselose Teilchen mit dem Spin $J = 1$. Sie sind durch die Quantenzahlen $B = Q = I = Y = C = 0$ gekennzeichnet. Man stellt sich vor, daß auf die Gluonen etwa die Hälfte des Nukleonimpulses entfällt, wodurch das in 3.2.5.3. erwähnte Impulsdefizit erklärt wird. Die Quantenchromodynamik ist die einzige Theorie der Hadronen, die gegenwärtig imstande ist, definitive Aussagen über die tief-inelastische Streuung zu machen; und diese sind mit den bisherigen experimentellen Ergebnissen im Einklang. Natürlich muß sie sich auch der Kernfrage nach der physikalischen Existenz von Quarks und Gluonen als beobachtbare Teilchen stellen. Hierzu gibt es die verschiedenartigsten Ansichten. Eine sowohl mathematisch als auch physikalisch denkbare Möglichkeit wäre, daß es zwar fundamentale Quark- und Gluonfelder gibt, diesen aber keine frei existie-

renden Teilchen entsprechen. Auch im Fall der Quantenchromodynamik können nur zukünftige Experimente und eine vertiefte theoretische Analyse erweisen, ob diese Theorie die Hadronen und die starke Wechselwirkung adäquat beschreibt. Man muß sich aber auch der Möglichkeit bewußt bleiben, daß neue Erfahrungsbereiche (höhere Energien) ganz neue physikalische Vorstellungen erforderlich machen könnten.

3.2.6.3. Weitere offene Fragen. Erwähnt sei das Problem der Zusammenfassung von schwacher, elektromagnetischer und starker Wechselwirkung in eine einheitliche Theorie, die als „*Große Unifikation*" bezeichnet wird. Sowohl die intermediären Bosonen (↑ 3.2.6.1.) als auch die Gluonen (↑ 3.2.6.2.) haben den Spin 1, sie werden daher als Mitglieder eines und desselben Supermultipletts betrachtet. Ebenso faßt man Leptonen und Quarks in Supermultipletts zusammen.

Eine interessante Konsequenz solcher einheitlichen Theorien ist, daß die Anzahl der Lepton- und Quarkarten sowie deren Quantenzahlen in bestimmten Verhältnissen zueinander stehen müssen. Von den Schwierigkeiten dieses Zugangs sei nur die Verletzung der Erhaltung der Baryonzahl genannt, die man durch geeignete Wahl offener Parameter in Einklang bringen muß mit der abgeschätzten Stabilität des Protons über mehr als 10^{30} Jahre hinweg.

Eine viel ältere Frage bezieht sich auf die Quantisierung der elektrischen Ladung: Weshalb treten in der Natur nur ganze Vielfache der elektrischen Elementarladung e (bzw. von $e/3$) auf? Im Unterschied zur Quantisierung des Spins oder Isospins liegt dem keine einfache mathematische Ursache zugrunde. Man weiß nur, daß eine Quantisierung der elektrischen Ladung resultiert, wenn magnetische Monopole, d. h. Teilchen, die sich wie ein isolierter Magnetpol verhalten, existieren. Trotz wiederholter Suche wurden solche Monopole bislang nicht gefunden, und das Problem ist noch ungelöst. Mag es sich bei alledem auch lediglich um theoretische Spekulationen handeln, so wird es doch ohne Zweifel die künftige Linie der Forschung beeinflussen.

3.3. Eigenschaften und Struktur der Atomkerne

3.3.1. Systematik und Stabilität der Kerne

3.3.1.1. Masse und Ladung. Diese beiden einfachsten Kenngrößen eines Kerns ergeben sich unmittelbar aus dessen Rolle innerhalb des Atoms (↑ 1.1.5.). Die Kernmasse stimmt nahezu mit der Atommasse überein; die wegen der bekannten Elektronenmasse genau angebbare Differenz stellt nur einen Anteil von 10^{-3} bis 10^{-4} der Gesamtmasse dar. Mit den Mitteln der Massenspektroskopie (↑ 1.1.3.5.) können so die Kernmassen sehr genau ermittelt werden. Sie sind bis auf Abweichungen der Ordnung 10^{-2} (s. a. 3.3.2.2.) ganzzahlige Vielfache AM_p der Masse des Wasserstoffkerns. A wird als *Massenzahl* bezeichnet.

Die elektrische Ladung $+Ze$ der Kerne ist positiv und gleich einem Vielfachen Z der Elementarladung (↑ 1.1.2.1.). Die *Kernladungszahl* Z ist die wichtigste Eingangsgröße für die elektronische Struktur des Atoms (↑ 4.2.1.2.) und kann daher aus einer Vielzahl atomarer Daten entnommen werden. So bestimmt Z als Ordnungszahl einerseits die Stellung des jeweiligen Elements im Periodensystem, andererseits über das Moseleysche Gesetz (↑ 1.1.4.7.) die Wellenlängen der charakteristischen Röntgenspektren.

Entsprechend der grundlegenden Strukturvorstellung (↑ 1.1.5.), daß alle Kerne aus *Nukleonen*, das ist die gemeinsame Bezeichnung für die etwa gleich schweren Neutronen und Protonen, aufgebaut sind, ist A die Gesamtzahl der im Kern enthaltenen Nukleonen. Z gibt die Zahl der Protonen an, die einfach positiv geladen sind, während die Neutronen, wie schon ihr Name andeutet, keine elektrische Ladung tragen. Kerne mit gleicher Ladungszahl Z, also mit gleicher Protonenzahl, werden *Isotope* genannt. Die zugehörigen Atome haben fast die gleiche Elektronenstruktur und damit auch gleiches chemisches Verhalten. Kerne mit gleicher Massenzahl A, also gleich viel Nukleonen, heißen *Isobare*; solche mit gleicher Neutronenzahl *Isotone*. Als Oberbegriff kann „Nuklid" gebraucht werden; dies ist gleichbedeutend mit „Kern", aber mit der Orientierung auf die Unterscheidung verschiedener Kerne.

Als *Symbol* zur Kennzeichnung eines Kerns benutzt man das chemische Symbol zusammen mit der Massenzahl, beispielsweise ist He-4, aus historischen Gründen oft als α-Teilchen bezeichnet, der aus zwei Neutronen und zwei Protonen bestehende Heliumkern; oder Ca-40 ist das Calciumisotop aus 20 Neutronen und 20 Protonen. Zur größeren Deutlichkeit kann die durch das chemische Symbol eindeutig fixierte Kernladungszahl zusätzlich angegeben werden, also $^{4}_{2}$He und $^{40}_{20}$Ca. Die meisten stabilen Isotope hat Zinn, und zwar insgesamt zehn von ^{112}Sn bis ^{124}Sn. Beispiele für isobare Kerne sind ^{40}Ar und ^{40}Ca. Ein Sonderfall isobarer Kerne sind die bei leichten Kernen anzutreffenden *Spiegelkerne*, die durch Vertauschung von Neutronen- und Protonenzahl auseinander hervorgehen, beispielsweise $^{14}_{8}$O und $^{14}_{6}$C.

3.3.1.2. Stabilität. Aufgrund verschiedener Umwandlungsmöglichkeiten ist ein Kern mit beliebiger Neutronen- und Protonenzahl im

allgemeinen instabil. Ein Kern kann seine Massenzahl A nur durch Aufnahme oder Abgabe von Nukleonen ändern. Der Vorgang der Aufnahme von Protonen oder anderen Kernen wird normalerweise durch die elektrostatische Abstoßung der Kerne verhindert. Bezüglich der Aufnahme von Neutronen wären nahezu alle Kerne instabil; freie Neutronen sind aber daher normalerweise nicht vorhanden. Vor allem im Kernreaktor führt die *Neutronenabsorption* zur *Bildung von Transuranen*, z. B. Neptunium, Plutonium. Die Abgabe einzelner Nukleonen ist wegen deren Bindung an den Kern nur aus hochangeregten Kernen oder aus bisher nicht hergestellten Kernen weitab von der Stabilitätslinie (↑ 3.3.1.3.) möglich. Dagegen ist es energetisch vorteilhafter und bei schwereren Kernen auch aus dem Grundzustand heraus möglich, fest gebundene Nukleonengruppen abzugeben. Dies geschieht bei der *Emission von α-Teilchen* und bei der *spontanen Kernspaltung* (↑ 3.4.1.).

Diese Prozesse begrenzen die Massenzahlen stabiler Kerne. U-238 ist der schwerste der in der Natur mit nennenswerter Häufigkeit vorkommenden Kerne. Jenseits des Bismut haben alle Kerne außer Th-232 und U-238 Halbwertszeiten für α-Zerfall, die kleiner als 10^9 Jahre sind. Die beiden genannten Kernarten sind nur wegen ihrer langen Lebensdauer seit ihrer Bildung noch nicht vollständig zerfallen. Die Instabilität der schwereren Transurane wird durch den Prozeß der spontanen Kernspaltung bestimmt. Die Spaltwahrscheinlichkeit nimmt mit wachsender Kernladung rapid zu. Der entscheidende Parameter ist Z^2/A (↑ 3.4.1.3.2.). Für $Z^2/A = 41$ erreicht die Lebensdauer den Sekundenbereich. Diese Grenze ist bei den bisher künstlich hergestellten Transuranen erreicht. Man hofft dennoch, im Massenbereich um $A \approx 300$ eine „Insel der Stabilität" mit relativ langlebigen Kernen zu finden.

Für Neutronen- und Protonenzahl einzeln gelten keine Erhaltungssätze, vielmehr erlauben die durch die schwache Wechselwirkung (↑ 3.2.4.5.) vermittelten Prozesse des β-Zerfalls, d. h. der Emission von Elektronen oder Positronen sowie des Elektroneneinfangs (↑ 3.4.1.2.2.), die Umwandlung von Neutronen in Protonen und umgekehrt, also Umwandlungsprozesse zwischen isobaren Kernen. Die schwache Wechselwirkung ist also, obwohl ohne Einfluß auf die Struktur der Kerne, dennoch sehr wesentlich für deren Stabilität.

Für die genauere Energiebilanz von Umwandlungen zwischen „benachbarten" isobaren Kernen muß die Massendifferenz $M_n - M_p \approx 2{,}5\ M_e$ von etwa 2,5 Elektronenmassen und die Masse des emittierten oder absorbierten Elektrons bzw. Positrons beachtet werden. Bei der Vergrößerung der Protonenzahl durch Emission eines Elektrons beim β^--Zerfall wird die Ruhenergie von $(2{,}5 - 1)\ M_e c^2$ frei. Bei dem hinsichtlich der Energiebilanz umgekehrten Prozeß der Vergrößerung der Neutronenzahl durch Einfang eines Elektrons muß die gleiche Energiemenge in Ruhenergie umgewandelt werden. Einer der beiden Prozesse ist also stets energetisch möglich, so daß von zwei benachbarten isobaren Kernen nur ein einziger stabil sein sollte (s. a. 3.3.2.2.4.).

Es hängt vom jeweils erreichten Stand der Experimentiertechnik ab, bis zu welch kleinen Lebensdauern man *instabile* Kerne nachweisen und identifizieren kann. Einige instabile Kerne treten als Zwischenprodukte in den Zerfallsreihen der natürlichen Radioaktivität (↑ 1.1.5.4.) auf. Die meisten werden jedoch künstlich bei Kernreaktionen (↑ 3.4.2.) aller Art erzeugt bzw. treten wiederum als Zwischenprodukte bei den Kernumwandlungen der primären Reaktionsprodukte auf. Durch den Einsatz künstlich hergestellter *Radionuklide* (↑ 3.4.3.) konnten zahlreiche wichtige Anwendungen erschlossen werden. Eine gewisse Sonderstellung haben die *Transurane*, also die Kerne mit $Z > 92$. Die ersten Transurane Neptunium ($Z = 93$) und Plutonium ($Z = 94$) können relativ leicht aus Uranium durch Neutronenabsorption und nachfolgenden β-Zerfall gebildet werden. Zur Erzielung höherer Massenzahlen beschießt man Uranium oder niedere Transurane, vor allem das ausreichend verfügbare Plutonium, mit energiereichen Ionen. Zunächst wurden solche Experimente mit α-Teilchen und leichten Ionen wie die von Kohlenstoff, Stickstoff und Sauerstoff durchgeführt, seit der Entwicklung von Schwerionenbeschleunigern auch mit schwereren Ionen (s. a. 3.4.2.5.5.).

3.3.1.3. Zusammensetzung der β-stabilen Kerne. Die Darstellung (Abb. 3.3.1-1) im N-Z-Diagramm zeigt deutlich, daß diese Kerne in einem schmalen Streifen um eine gedachte *Stabilitätslinie* liegen. Bei gegebener Nukleonenzahl gibt es offenbar eine energetisch besonders günstige Aufteilung auf Neutronen und Protonen. Kerne in einigem Abstand von der Stabilitätslinie sind

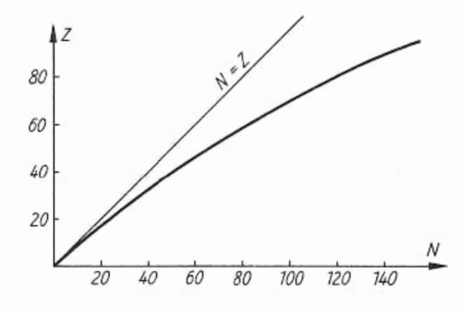


Abb. 3.3.1-1 Zusammensetzung der β-stabilen Kerne

stets instabil, mit wachsender Abweichung von dieser Linie werden die Lebensdauern so klein, daß sich solche Kerne dem Nachweis entziehen. Die stabilen leichten Kerne haben etwa gleich viel Neutronen wie Protonen, während sich bei mittleren und schweren Kernen ein mit der Massenzahl anwachsender *Neutronenüberschuß* herausbildet.

Eine genauere Untersuchung zeigt eine deutliche *Bevorzugung gerader Nukleonenzahlen.* In der Reihenfolge (N, Z) und mit den Symbolen g bzw. u für gerades bzw. ungerades N oder Z gibt es 164 stabile (g,g), 50 (g,u), 55 (u,g) und nur 4 (u,u)-Kerne, nämlich $^2_1$H, also das Deuteron, $^6_3$Li, $^{10}_5$B und $^{14}_7$N. Außerdem gelten die Mattauchschen *Isobarenregeln.* Für ungerades $A = N + Z$ existiert jeweils nur ein einziges stabiles Nuklid, also entweder ein (g,u)- oder (u,g)-Kern, während zu geradem A im allgemeinen mehrere stabile Isobare gehören; bis auf die erwähnten Ausnahmen sind das (g,g)-Kerne, wobei sich die Protonen- und Neutronenzahlen benachbarter stabiler Isobare um 2 unterscheiden.

3.3.1.4. Isotopieeffekte und Isotopentrennung. Kernphysikalisch sind isotope Kerne wesentlich verschieden, d. h. ihre Energie, ihr Anregungsspektrum, ihre Zerfallsmöglichkeiten und -wahrscheinlichkeiten unterscheiden sich deutlich voneinander. Im nicht-kernphysikalischen Bereich treten dagegen im allgemeinen nur kleine Effekte auf, die von den Unterschieden in Kernmasse und Kernvolumen und von den schwachen magnetischen Effekten der Kerne bewirkt werden. Dazu zählen die Isotopieverschiebungen und die Hyperfeinstruktur (↑ 4.2.4.8.) von *Atom- und Molekülspektren.* Die ausgeprägtesten Isotopieeffekte zeigen sich bei der *Ablenkung von Ionen* in elektrischen und magnetischen Feldern (↑ 1.1.2.), da hier die Ablenkungen unmittelbar durch die Teilchenmasse bestimmt, nämlich umgekehrt proportional zu ihr sind. Die auf dieser Grundlage entwickelte moderne Massenspektroskopie gestattet eine völlige Isotopentrennung und die genaue Bestimmung der relativen Anteile der einzelnen Isotope. Ebenfalls direkt macht sich die unterschiedliche Masse der Isotope in rasch rotierenden Zentrifugen, sogenannten *Ultrazentrifugen,* bemerkbar. Durch die starken Zentrifugalkräfte werden die schwereren Isotope außen angereichert. Der Massenunterschied der Isotope führt auch zu physikalischen und chemischen Effekten bei *Phasen- und Stoffumwandlungen.* Meist reichern sich dabei die leichteren Isotope in der flüchtigeren Phase bzw. in den kleineren Molekülen an. Diese Effekte sind, von wenigen Ausnahmen abgesehen, sehr klein, führen aber auch in der Natur zu beobachtbaren Verschiebungen der Isotopenzusammensetzung in Abhängigkeit von der Vorgeschichte und dem Fundort der Probe. Beispielsweise sind im Regen- und Grundwasser die leichten Nuklide H-1 und O-16, aber im Meerwasser die schweren H-2 und O-18 angereichert. Anreicherungs- bzw. Verarmungsprozesse wurden im Laufe der Jahrmillionen auch durch Einbeziehung der Elemente in den Biozyklus bewirkt.

Der Massenunterschied zeigt sich nicht nur im physikalisch-chemischen Gleichgewicht, sondern auch in der Geschwindigkeit, mit der Nichtgleichgewichtsprozesse ablaufen. So diffundieren die Moleküle mit den leichten Isotopen schneller durch poröse Wände als die schweren, und bei der elektrolytischen Zerlegung von Wasser werden die Moleküle mit schwerem Wasserstoff im Rückstand angereichert. Besondere Bedeutung zur Isotopentrennung hat die *Thermodiffusion* erlangt, das ist die durch Temperaturunterschiede in einem Gasgemisch hervorgerufene Diffusion.

Für Forschungszwecke und vor allem für die Kernenergetik (↑ 3.5.) werden möglichst isotopenreine oder wenigstens mit bestimmten Isotopen angereicherte Stoffe benötigt. Eine Isotopentrennung in einem Schritt, aber bei größeren Substanzmengen mit sehr hohem Aufwand, gelingt durch *elektromagnetische Massentrenner,* die nach dem Prinzip der Massenspektrometer arbeiten (Tafel 24). Neuerdings wurde es durch den *Einsatz leistungsstarker Laser* mit genau abgestimmter Frequenz möglich, die sehr kleinen Isotopieeffekte in den atomaren Energieniveaus zur Isotopentrennung zu nutzen. Beispielsweise kann die Frequenz des eingestrahlten Laserlichts so gewählt werden, daß nur Atome des einen Isotops ionisiert werden und demzufolge durch ein angelegtes elektrisches Feld aus der Mischung herausgeführt werden können.

Wegen der Kleinheit der meisten Isotopieeffekte muß der Trennungsvorgang in vielen aufeinanderfolgenden Stufen wiederholt werden. Durch entsprechende Diffusions- oder Destillationskolonnen wird der Trennfaktor der Einzelstufe potenziert. Wird beispielsweise der Anteil eines anzureichernden Isotops beim Durchlaufen eines Trennrohrs um den Faktor $a > 1$ vergrößert, so bringt eine Kolonne aus N Rohren eine Anreicherung um den Faktor a^N (Tafel 17).

3.3.2. Meßbare Kerngrößen

3.3.2.1. Größe und Form. Bereits aus den frühen Experimenten von Rutherford u. a. zur Streuung von α-Teilchen wurde geschlossen, daß die Kernvolumina etwa proportional zur Massenzahl A, die *Kernradien* also proportional zu $A^{1/3}$ sind. Genauere experimentelle Untersuchungen haben dies bestätigt und präzisiert.

Dabei beruhen die genauesten Verfahren auf der gut bekannten elektrischen Wechselwirkung mit dem Kern und messen daher streng genommen nur dessen Ladungsverteilung, also die Verteilung der Protonen. Ein wesentlicher Unterschied zwischen Protonen- und Neutronenverteilung liegt jedoch sicher nicht vor. Als für die Größe der Kerne charakteristische Länge erweist sich 10^{-15} m = 1 Fermi = 1 fm; diese Einheit wird bei kernphysikalischen Längenangaben benutzt.

3.3.2.1.1. *Untersuchungsmethoden.* Die endliche Größe des Kerns macht sich in den Atomspektren durch eine Isotopieverschiebung (↑ 4.2.4.9.) bemerkbar. Dieser Effekt ist wegen der geringen Elektronendichte am Kernort sehr klein, aber wegen der hohen spektroskopischen Genauigkeit gut meßbar. Größere Verschiebungen gegenüber dem idealisierten Modell eines punktförmigen Kerns treten bei Myonenatomen (↑ 4.2.2.5.) auf, die daher geeignete Objekte für das Studium der Kerngröße sind.

Als sehr aussagekräftig erwiesen sich die Versuche zur *Streuung schneller Elektronen* an Kernen. Um dabei Einzelheiten der Ladungsverteilung nachweisen zu können, muß die Elektronenwellenlänge (↑ 1.2.2.3.) von der Größenordnung der Kernradien oder kleiner sein (↑ 2.3.2.4.). Das ist erfüllt bei Elektronen mit Energien von einigen Hundert MeV, die sich also fast mit Lichtgeschwindigkeit bewegen.

Weniger genaue Methoden beruhen auf der Streuung von Neutronen, Protonen und α-Teilchen sowie auf den allein durch die elektrostatische Protonenwechselwirkung bedingten Energiedifferenzen zwischen Spiegelkernen (↑ 3.3.1.1.).

3.3.2.1.2. *Ergebnisse.* Die meisten Kerne sind in guter Näherung kugelförmig. Die *Dichte im Zentrum* ist unabhängig von der Massenzahl, der *Radius* ist $R = r_0 A^{1/3}$ mit $r_0 \approx 1$ fm. R ist jedoch ohne nähere Angaben nicht genau definiert, da die Nukleonendichte am Kernrand in einem ebenfalls A-unabhängigen Bereich der Dicke 2,4 fm von 90% auf 10% ihres Innenwertes abfällt. Definiert man den Radius $R_{1/2}$ so, daß dort die Dichte auf die Hälfte gesunken ist, so erhält man für den Parameter r_0 etwa 1,1 fm. Die gemessenen Verteilungen (Abb. 3.3.2-1)

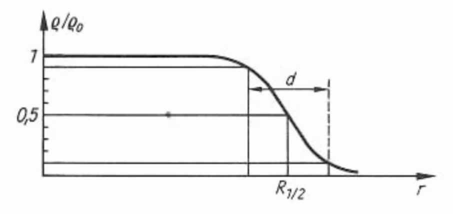


Abb. 3.3.2-1 Nukleonenverteilung im Kern (Erläuterung im Text)

können sehr gut für alle $A > 16$ durch $\varrho = \varrho_0/[1 + \exp(r - R_{1/2}/d)]$ mit $\varrho_0 = 0{,}17$ Nukleonen/fm³ und $d = 0{,}54$ fm dargestellt werden. Der Wert von $R_{1/2}$ ergibt sich daraus, daß das Volumenintegral über ϱ die Nukleonenzahl A liefern muß; genähert gilt die obige Proportionalität zu $A^{1/3}$.

3.3.2.1.3. *Deformierte Kerne.* Kerne mit Massenzahlen $150 < A < 190$ und $A > 222$ zeigen zum Teil starke Abweichungen von der Kugelform. Sie haben die Gestalt von Rotationsellipsoiden, beispielsweise wurde für Cd-158 ein Hauptachsenverhältnis $b/a \approx 0{,}7$ abgeschätzt, a ist längs der Rotationsachse gemessen. Experimentelle Daten, aus denen man auf die Kerndeformationen schließen kann, betreffen anormal große Isotopieverschiebungen, den als *Coulombanregung* bezeichneten Prozeß, bei dem ein deformierter Kern ohne direkten „Kontakt" über Kernwechselwirkungen durch ein vorbeifliegendes geladenes Teilchen angeregt wird, und die Rotationszustände (↑ 3.3.3.4.) dieser Kerne. Bei den Isotopieverschiebungen wird nicht die volle Abweichung von der Kugelsymmetrie, sondern nur der Mittelwert über die schwankende Kernorientierung wirksam. Kerne mit dem Spin $I = 0$ oder $I = 1/2$ wirken auf die umgebenden Elektronen generell wie kugelförmige Kerne, unabhängig von ihrer eigentlichen Gestalt.

3.3.2.2. Energie der Kerne

3.3.2.2.1. *Massendefekt und Bindungsenergie.* Die massenspektroskopisch sehr genau bestimmbaren Kern- bzw. Atommassen werden gewöhnlich in der physikalischen Atommassenskala durch dimensionslose Zahlen ausgedrückt. Die *atomare Masseeinheit* wird definiert durch die Festlegung: Masse des Kohlenstoffatoms C-12 gleich 12, 000 000 u; 1 u = $1{,}660277 \cdot 10^{-27}$ kg. Die Masse $M(N,Z)$ eines Kerns aus N Neutronen und Z Protonen ist kleiner als die Summe der Massen seiner Bestandteile, die Differenz $\Delta M = NM_n + ZM_p - M(N,Z)$ heißt Massendefekt.

Der Massendefekt liefert aufgrund der Einsteinschen Energie-Masse-Beziehung (↑ 2.1.3.6.) die *Bindungsenergie* $E_B = \Delta M/c^2$ des Kerns, das ist also die Energie, die man zuführen müßte, um den Kern in seine Bestandteile zu zerlegen. – $\Delta M/c^2$ ist die Energie des Kerns ohne Berücksichtigung der Ruhenergien der Nukleonen, also diejenige Energie, die durch die Bewegung und gegenseitge Wechselwirkung der Nukleonen im Kern zustande kommt. Zur Umrechnung von Massen in Energien dient die Äquivalenz 1 u = 931,4 MeV.

Die Größenordnung von einigen MeV Bindungsenergie pro Nukleon, verglichen mit den Ionisierungsenergien von einigen eV in der Atomhülle, verdeutlicht, daß die Kerne außerordentlich fest gebundene Einheiten darstellen (s. a.

3.3.4.), die bei den normalerweise erreichbaren Temperaturen unbeeinflußt bleiben und nur durch „innere" Instabilität (↑ 3.3.1.2.) oder durch Beschuß mit hochenergetischen Teilchen zu Umwandlungen veranlaßt werden. Gelingt es aber, wie bei der Kettenreaktion der Spaltung oder der thermisch induzierten Kernfusion, makroskopische Mengen von Kernen zu Reaktionen zu veranlassen, bei denen Energie freigesetzt wird, so entstehen Energiequellen neuer Dimension (↑ 3.5., ↑ 9.4.).

Die Energie bzw. Bindungsenergie der stabilen Kerne ist, abgesehen von den allerleichtesten Kernen, genähert proportional zur Massenzahl A mit $E_B/A \approx 8$ MeV. Eine genauere Analyse (Abb. 3.3.2-2) zeigt ein flaches Minimum der

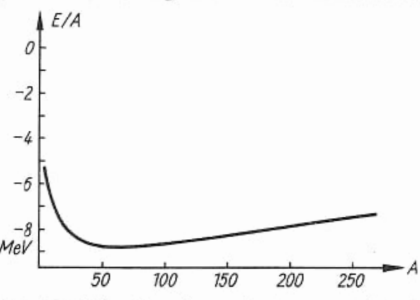


Abb. 3.3.2-2 Energie pro Nukleon als Funktion der Massenzahl

Energie pro Nukleon im Bereich von $A = 60$ bis 65 und ein allmähliches Ansteigen dieser Energie bei schwereren Kernen. Bei leichten Kernen wirken vor allem Oberflächeneffekte, bei schwereren die elektrostatische Abstoßung zwischen den Protonen bindungslockernd (↑ 3.3.2.2.2.). Sowohl durch Verschmelzung (Fusion) leichter als auch durch Spaltung schwerer Kerne kann die Energie pro Nukleon erniedrigt, also Energie freigesetzt werden.

3.3.2.2.2. *Bethe-Weizsäcker-Formel.* Die gemessenen Kernenergien können mit einer relativen Genauigkeit von etwa 10^{-2} durch die Gleichung

$$(1) \quad E(N,Z) = -aA + bA^{2/3} + c\left(\frac{N-Z}{N+Z}\right)^2 A + d\frac{Z^2}{A^{1/3}}$$

dargestellt werden. Ein möglicher Konstantensatz, bei dem der Wert von d in guter Übereinstimmung mit den experimentell ermittelten Ladungsverteilungen (↑ 3.3.2.1.) im Kern ist, lautet $a = 15{,}75$ MeV, $b = 17{,}8$ MeV, $c = 23{,}7$ MeV, $d = 0{,}710$ MeV. Formel 3.3.(1) ist mehr als eine empirische Interpolationsformel, da ihre einzelnen Terme physikalisch interpretierbar sind. Dabei ist zu beachten, da dßas Kernvolumen proportional zu A, der Kernradius (↑ 3.3.2.1.) also $\sim A^{1/3}$ ist. Der erste und der dritte Term von $E(N,Z)$ sind proportional zu A, sie sind also *Volumenterme*. Könnte man die elektrostatische Wechselwirkung zwischen den Protonen „ausschalten", was gedanklich möglich und für die Struktur des Einzelkerns nahezu bedeutungslos wäre, so wären Kerne mit beliebig großen Massenzahlen stabil. Dem entspricht das Konzept einer hypothetischen *Kernmaterie* aus sehr vielen Nukleonen; $-a + c(N-Z)^2/(N+Z)^2$ wäre die Energie pro Nukleon in dieser ausgedehnten Kernmaterie. Diese Energie pro Teilchen ist also, so wie in einem beliebigen Stück Materie, unabhängig von der Gesamtzahl der Teilchen, wohl aber muß sie aufgrund der deutlichen Bevorzugung gleicher Protonen- und Neutronenzahlen (↑ 3.3.1.3.) von dem *relativen* Neutronen- bzw. Protonenanteil abhängen. Für $N = Z$ hat sie ihren minimalen Wert, was durch die zu $(N-Z)^2/(N+Z)^2$ proportionale *Symmetrieenergie* ausgedrückt wird.

Der Term $bA^{2/3}$ in 3.3.(1) ist proportional zum Quadrat des Kernradius. Er erfaßt eine Bindungslockerung durch Oberflächeneffekte, die besonders anschaulich durch eine Analogie zwischen den Nukleonen im Kern und den Molekülen in einem Flüssigkeitstropfen werden. Der *Oberflächeneffekt* ist aber nicht an diese Analogie gebunden, sondern kann auch im Modell einer nahezu freien Bewegung der Nukleonen im Kern erklärt werden.

Der Term $dZ^2/A^{1/3}$ repräsentiert schließlich die Energie der *elektrostatischen Protonenabstoßung*. Eine homogen geladene Kugel hat nach den Gesetzen der Elektrodynamik eine elektrische Feldenergie (↑ 2.2.3.7.) der Größe $(3/5)\ Q^2/\pi 4\varepsilon_0 R$, dabei ist Q die Ladung und R der Radius der Kugel. Mit $Q = Ze$ und $R = r_0 A^{1/3}$ folgt so $d = (3/5)\ e^2/4\pi\varepsilon_0 r_0$.

Die Bethe-Weizsäcker-Formel berücksichtigt in der angegebenen Form weder die Besonderheiten der leichtesten Kerne, noch die Bevorzugung gerader Nukleonenzahlen, noch Anomalien in der Umgebung magischer Nukleonenzahlen (↑ 3.3.3.1.).

3.3.2.2.3. *Auswertung der Bethe-Weizsäcker-Formel.* Obwohl die eingehenden Parameter a bis d durch Anpassung an die gemessenen Energien der stabilen bzw. langlebigen Kerne bestimmt werden können, ist die Formel auch für instabile Kerne anwendbar. Der *Verlauf der Stabilitätslinie* wird bestimmt durch die Forderung, daß E bei festem A, also beim Vergleich isobarer Kerne, sein Minimum annimmt. Längs eines Isobarenschnitts $N + Z = A = \text{const}$ ändern sich die Symmetrieenergie und die elektrostatische Energie. Das Resultat $(N/Z)_{\text{stab}} = 1 + (d/2c)A^{2/3}$ zeigt deutlich die beobachtete Bevorzugung neutronenreicher Kerne (↑ Abb. 3.3.2-1). Die *Energie längs der Stabilitätslinie*

$E_{stab} = -aA + bA^{2/3} + cA[1 + 4c/dA^{2/3}]^{-1}$ gibt die Bindungslockerung durch den Oberflächen- und durch einen minimalisierten Symmetrie-Coulomb-Term wieder. Damit wird das flache Minimum von E_{stab}/A bei mittleren Massenzahlen erklärt (↑ Abb. 3.3.2-2). Insgesamt wird die Energie gegenüber dem Volumenterm $-aA$ auf etwa die Hälfte verringert.

Der Neutronenüberschuß ist eine Folge der elektrostatischen Abstoßung zwischen den Protonen und daher proportional zu d. Für $d = 0$ wäre $(N/Z)_{stab} = 1$ und $E_{stab} = -aA + bA^{2/3}$. Dann wären auch beliebig schwere Kerne stabil gegen α-Zerfall und Spaltung. Die gegenüber den eigentlichen Kernkräften sehr schwachen elektrischen Kräfte führen zur Instabilität schwerer Kerne.

Außerhalb der Stabilitätslinie steigt die Energie längs eines jeden Isobarenschnitts an (Abb. 3.3.2-3 und 3.3.2-4a), und zwar bei kleinen Massenzahlen, wo eine Änderung von N und Z um $+1$ bzw. -1 eine wesentliche Veränderung des „Mischungsverhältnisses" bedeutet, wesentlich steiler als bei größerem A.

3.3.2.2.4. *Paarungsenergie und Isobarenregeln.* Die Bevorzugung gerader Nukleonenzahlen (↑ 3.3.1.3.) beruht auf einer niedrigeren Energie, also einer festeren Bindung der entsprechenden Kerne. Dies kann durch die sogenannte Paarungsenergie

$$E_p = \begin{cases} +\Delta \text{ für (u,u)-Kerne} \\ 0 \text{ für (u,g)- und (g,u)-Kerne} \\ -\Delta \text{ für (g,g)-Kerne} \end{cases}$$

beschrieben werden, wobei die aus den gemessenen Energien entnommenen Δ-Werte im Mittel durch $\Delta = 12\ \text{MeV}/\sqrt{A}$ darstellbar sind. Der ausgeprägte Paarungseffekt kann als Bildung von gebundenen Neutronen- bzw. Protonenpaaren interpretiert werden, in Analogie zu den Cooperpaaren bei der Supraleitung (↑ 8.5.3.2.).

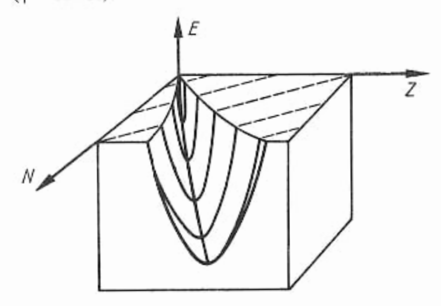


Abb. 3.3.2-3 Energie der Kerne als Funktion von Neutronen- und Protonenzahl

Für Kerne mit ungerader Massenzahl ist $E_p = 0$. Daher ist von allen isobaren Kernen nur einer stabil (↑ Abb. 3.3.2-4a), in Einklang mit der entsprechenden empirischen Isobarenregel (↑ 3.3.1.3.). Dagegen spaltet bei Kernen mit gerader Massenzahl die Energiefläche aufgrund der Paarungsenergie gewissermaßen in zwei Flächen auf, eine für die (g,g)- und eine für die (u,u)-Kerne (↑ Abb. 3.3.2-4b und c). Die durch β-Prozesse auseinander hervorgehenden Kerne liegen abwechselnd auf der oberen und unteren Kurve. Daher gibt es im allgemeinen keine stabilen (u,u)-, aber mehrere stabile (g,g)-Kerne (↑ 3.3.1.3.). Die (g,g)-Kerne höherer Energie könnten sich prinzipiell in die energieärmeren durch Emission bzw. Einfang zweier Elektronen umwandeln. Diese Prozesse sind aber so un-

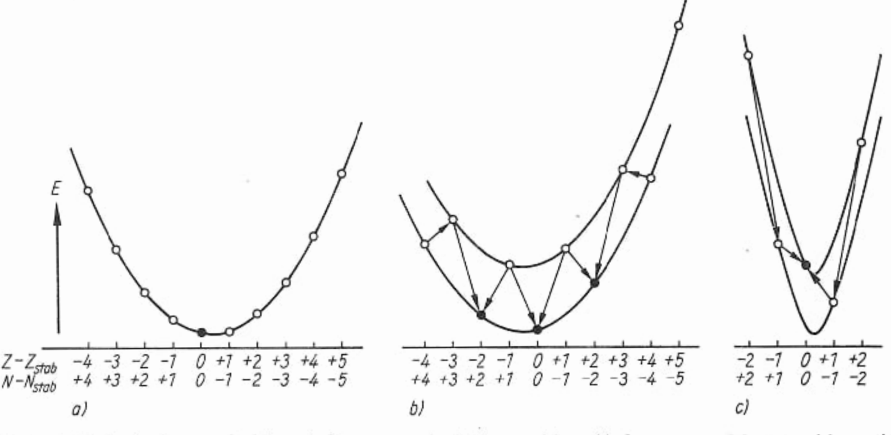


Abb. 3.3.2-4 Isobarenschnitte; a) für ungerade Massenzahlen, b) für gerade Massenzahlen mit Berücksichtigung der Paarungsenergie für mittlere oder schwere Kerne, c) für gerade Massenzahlen mit Berücksichtigung der Paarungsenergie für leichte Kerne

wahrscheinlich, daß die genannten Kerne praktisch stabil sind. Die Stabilität einiger leichter (u,u)-Kerne beruht auf dem steilen Anstieg der Energiekurven bei kleinen A-Werten (↑ Abb. 3.3.2-4c).

3.3.2.3. Angeregte Kernzustände

3.3.2.3.1. *Termschemata.* So wie die meisten gebundenen quantenmechanischen Systeme besitzen auch die Kerne Anregungszustände mit diskreten Energien. Jeder Kern besitzt ein für ihn typisches „individuelles“ Anregungsspektrum, das aus sehr vielen, mit wachsender Anregungsenergie immer dichter liegenden Niveaus besteht. Die systematische Erforschung dieser Anregungszustände liefert daher ein riesiges experimentelles Material. Eine wichtige Untersuchungsmethode ist die *Anregung von Kernen* durch inelastische Neutronenstreuung, durch vorbeifliegende geladene Teilchen oder durch Absorption energiereicher elektromagnetischer Strahlung. Dabei kann durch Messung des Energieverlusts des anregenden Teilchens bzw. der Photonenenergie eine direkte Aussage über die Energie des erzeugten Anregungszustands gemacht werden, vor allem aber können die Folgeprozesse systematisch untersucht werden. Dies sind häufig Kaskaden von γ-Prozessen (↑ 3.4.1.1.), kombiniert mit β-Zerfällen (↑ 3.3.1.2.). Daher betreffen die auf diese Weise gewonnenen Termschemata (Abb. 3.3.2-5) nicht nur einen bestimmten Kern, sondern eine Gruppe isobarer Kerne. Zur eindeutigen Zuordnung

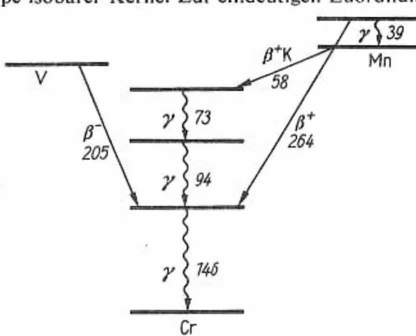


Abb. 3.3.2-5 Termschema mit β- und γ-Übergängen; die Zahlen bedeuten gemessene Energiedifferenzen in keV

der Übergänge müssen sorgfältige Koinzidenzmessungen durchgeführt werden, um festzustellen, welche γ-Übergänge (schnell) an welche β-Übergänge (langsam) anschließen. Der Ausgangszustand einer Zerfallskaskade kann auch ein durch eine beliebige Kernreaktion hergestellter instabiler Kern sein.

Eine weitere Untersuchungsmöglichkeit ist die Messung der *Energieabhängigkeit des Wirkungsquerschnitts* (↑ 2.4.3.5.) von durch Neutronen ausgelösten Kernreaktionen (↑ 3.4.2.5.1.). Die ausgeprägten Resonanzen im Verlauf des Wirkungsquerschnitts entsprechen den Energien des nach Absorption des Neutrons gebildeten Compoundkerns (↑ 3.4.2.3.). Da dieser sogar bei Absorption eines Neutrons *ohne* kinetische Energie dessen Bindungsenergie von etwa 8 MeV als Anregungsenergie erhält, können mit dieser Methode nur angeregte Zustände oberhalb dieser Anregungsenergie untersucht werden.

Außerdem ist der Nachweis von Teilchengruppen unterschiedlicher Energie beim α-Zerfall oder bei den leichten *Endteilchen einer Kernreaktion* möglich. Die Energiedifferenzen entsprechen den Anregungsenergien des gebildeten Endkerns, beispielsweise konnten bei der Reaktion B-10 (αp) C-13 bei einheitlicher Energie der α-Teilchen vier verschiedene Protonenenergien nachgewiesen werden.

Der erste angeregte Zustand leichter Kerne liegt meist 1 bis 2 MeV über dem Grundzustand, extrem hoch bei O-16 mit 6,05 MeV. Bei schwereren Kernen liegen die unteren Anregungszustände meist im 100-keV-Bereich. Das gilt besonders für die Rotationsanregungen deformierter Kerne (↑ 3.3.3.4.).

3.3.2.3.2. *Zerfallswahrscheinlichkeiten und Niveaubreiten.* Die angeregten Zustände, bei instabilen Kernen auch der Grundzustand, können durch ihre Zerfallswahrscheinlichkeiten γ (↑ 2.4.3.4.) charakterisiert werden. Die mittlere Lebensdauer ist $\tau = \gamma^{-1}$, die oft benutzte Halbwertszeit $T_{1/2} = \tau \ln 2$. Die Zerfallswahrscheinlichkeit schließt alle von dem betrachteten Zustand aus möglichen Übergänge ein, denen partielle Übergangswahrscheinlichkeiten γ_i zugeordnet werden. Es gilt $\gamma = \sum_i \gamma_i$. Die Verhältnisse der verschiedenen γ_i ergeben die Verzweigungsverhältnisse für die verschiedenen Zerfallsmöglichkeiten.

Bei hoch angeregten Kernen, Anregungsenergie > 8 MeV, wird die Emission von Nukleonen möglich, insbesondere die von Neutronen, die im Unterschied zur Aussendung geladener Teilchen nicht durch den Coulombwall (↑ 3.4.1.3.1.) behindert wird. Ein typischer Wert ist $\gamma_n \approx 10^{15}\,\mathrm{s}^{-1}$. Die Lebensdauern von sich vorwiegend durch Emission von γ-Quanten umwandelnden Zuständen variieren etwa zwischen 10^{-10} und 10^{-17} s. Ausnahmsweise können diese Lebensdauern aber auch den Bereich von Sekunden oder Tagen erreichen, nämlich dann, wenn beim Übergang große Spindifferenzen (↑ 3.3.2.4.) zwischen dem angeregten Ausgangszustand und allen tieferliegenden Zuständen überbrückt werden müssen. Solche langlebigen Anregungszustände werden isomere Zustände genannt; *Isomere* sind gleiche Kerne mit verschiedener Energie. Außer bei Isomerie können

α- und β-Prozesse nicht mit Strahlungsübergängen konkurrieren.
Die endliche Lebensdauer eines Zustands führt zu einer Unschärfe ΔE seiner Energie, er ist gewissermaßen auf ein Intervall dieser Breite verschmiert (↑ 2.5.6.3.). Es gilt $\tau \Delta E \approx \hbar$, also $\Delta E \approx \hbar\gamma$. Diese Niveaubreiten ΔE werden in der Kernphysik mit Γ bezeichnet. Entsprechend dem additiven Verhalten $\gamma = \sum_i \gamma_i$ der Übergangswahrscheinlichkeiten setzt sich auch die Breite $\Gamma = \sum_i \Gamma_i$ eines Niveaus aus den Partialbreiten $\Gamma_i = \hbar\gamma_i$ zusammen, die den einzelnen Prozessen zuzuordnen sind. Zu $\gamma \approx 10^{15}\,\mathrm{s}^{-1}$ gehört $\Gamma \approx 1\,\mathrm{eV}$, die Breiten sind daher auch bei mittleren Anregungsenergien noch klein gegen die Niveauabstände. Die Energieunschärfe eines beliebigen Übergangs, etwa die der Photonenenergie bei einem Strahlungsübergang, ist gleich der Summe der Breiten der beiden beteiligten Niveaus, unabhängig davon, von welchen Prozessen diese Breiten verursacht werden.

3.3.2.4. Spin und magnetisches Moment. Für den Eigendrehimpuls oder Spin (↑ 2.1.3.9.) der Kerne gelten alle allgemeinen Aussagen über den Drehimpuls quantenmechanischer Systeme (↑ 2.5.3.). Dementsprechend hat das Quadrat des Kerndrehimpulses nur die möglichen Werte $\hbar^2 I(I+1)$, wobei die dimensionslose Zahl I kurz als *Kernspin* bezeichnet wird. Da Nukleonen die Spinquantenzahl $s = 1/2$ haben und die Bahndrehimpulse der Nukleonen durch ganzzahlige Quantenzahlen charakterisiert sind, ist nach den quantenmechanischen Regeln für die Addition von Drehimpulsen (↑ 2.5.3.4.) der Kernspin ganzzahlig ($I = 0, 1, 2 \ldots$) für Kerne mit gerader Massenzahl A und halbzahlig ($I = 1/2, 3/2, \ldots$) für ungerade A. Da der Kernspin nicht direkt gemessen werden kann, müssen indirekte Methoden zu seiner Bestimmung benutzt werden, die zum Teil ohne eingehende theoretische Analyse nicht verständlich sind. Am einfachsten, aber nur für stabile oder metastabile Kernzustände anwendbar, ist die Ermittlung der Zahl $2I+1$ der verschiedenen *Orientierungsmöglichkeiten* des mit dem Kernspin verbundenen magnetischen Moments. Mit der gleichen Beschränkung können charakteristische *Intensitätswechsel in den Rotationsspektren* (↑ 5.3.1.1.) von aus zwei gleichen Atomen (auch gleiche Isotope!) aufgebauten Molekülen zur Bestimmung von I dienen. Dieser Effekt ist unabhängig vom magnetischen Kernmoment; er wird durch die Forderung nach Antisymmetrie der Molekülwellenfunktion (↑ 2.5.4.) gegenüber Vertauschung der beiden Kerne bewirkt. Aufeinanderfolgende Linien einer Rotationsbande zeigen ein Intensitätsverhältnis von $(I+1):I$. Für $I = 0$ verschwindet im Rotationsspektrum jede zweite Linie. Zur Bestimmung der Spins kurzlebiger angeregter Kernzustände sind nur rein kernphysikalische Methoden brauchbar, wichtig ist die Bestimmung der Winkelverteilung von beim Zerfall emittierten Strahlungen sowie von bei Kernreaktionen entstehenden Teilchen.
Das mit dem Kernspin verbundene magnetische Moment hat die Größenordnung eines *Kernmagnetons* $e\hbar/2M_p$, das rund um den Faktor 2000 kleiner ist als das die Momente der Atomhülle charakterisierende Bohrsche Magneton. Die g-Faktoren (↑ 2.2.2.5.) der Nukleonen haben die Werte $g_p \approx 5{,}58$ für das Proton und $g_n \approx -3{,}82$ für das Neutron. Das Neutron besitzt trotz verschwindender Ladung ein magnetisches Moment, was als Hinweis auf eine innere Struktur gedeutet werden kann (s. a. 3.2.5.). Zur Bestimmung der Kernmomente stabiler Kerne, und zusammen mit ihnen der Kernspins, können die magnetischen Wechselwirkungseffekte zwischen Kern und Hülle dienen, die zur *Hyperfeinstruktur* (↑ 4.2.4.8.) der Atom- und Molekülspektren führen. Wesentlich direkter und genauer ist die Methode der kernmagnetischen Resonanz (↑ 4.4.3.).
Alle (g,g)-Kerne haben im Grundzustand keinen Spin, also $I = 0$, und daher auch kein magnetisches Moment.

3.3.3. Kernmodelle

Modellvorstellungen über die innere Struktur der Kerne spielen eine größere Rolle als Modelle für den Atombau, weil wegen der nicht genau bekannten und ziemlich komplizierten Nukleon-Nukleon-Wechselwirkungen (↑ 3.3.4.) die Theorie der Kernstruktur nicht auf einer so sicheren Grundlage wie die Hüllenphysik aufbauen kann (↑ 4.2.1.2.). In der Anfangszeit der Kernphysik dominierte mit dem *Tröpfchenmodell* für die energetischen Verhältnisse und dem *Sandsackmodell* für das Anfangsstadium einer Kernreaktion die Vorstellung, daß jedes Nukleon ständig mit vielen anderen in einer intensiven Wechselwirkung stehe, also auch Energie mit diesen anderen austausche und somit seinen eigenen Zustand laufend ändere. Inzwischen hat sich jedoch gezeigt, daß auch die Nukleonen im Kern in guter Näherung durch ein *Modell unabhängiger Teilchen* (↑ 2.5.5.1.) beschrieben werden können.

3.3.3.1. Schalenabschlüsse beim Kernbau. Die unabhängige Elektronenbewegung in der Hülle findet ihren prägnanten Ausdruck in der Elektronenkonfiguration (↑ 4.2.3.3.) des jeweiligen Atoms, auf deren Grundlage man zur theoretischen Deutung des Periodensystems (↑ 4.2.3.4.) gelangt. Umgekehrt ist der nichtmonotone Verlauf der Atomeigenschaften, ist die Existenz ausgezeichneter Elektronenzahlen, die den ab-

geschlossenen Schalen der Edelgase entsprechen, ein deutlicher Hinweis auf die relativ freie Bewegung der Elektronen. Denn in einem „Tröpfchenbild" stark wechselwirkender Teilchen vermögen sich diese wegen ihrer ungenügenden Beweglichkeit „nicht zu zählen", also gibt es keine besonders ausgezeichneten Teilchenzahlen. Zahlreiche Daten der Kernphysik beweisen die Existenz ausgezeichneter Nukleonenzahlen, in deren Umgebung charakteristische Nicht-Monotonien auftreten. Da diese Effekte von den zunächst herrschenden Strukturvorstellungen her unverständlich waren, bezeichnete man diese Zahlen als *magische Zahlen*. Sie lauten 2, 8, 20, 28, 50, 82, 126. Außer der nur als Neutronenzahl realisierten 126 können alle anderen sowohl als Neutronen- wie als Protonenzahlen auftreten. Besonders stabil sind sogenannte doppelt magische Kerne wie das α-Teilchen He-4 mit $\{N = 2, Z = 2\}$, das mit einer Bindungsenergie von reichlich 7 MeV/Nukleon deutlich über seinen Nachbarkernen liegt und schon fast den (E/A)-Wert schwererer Kerne erreicht; oder wie der O-16-Kern mit $\{N = 8, Z = 8\}$, dessen große Bindungsenergie z. B. beim Vergleich der Anlagerungsreaktionen C-12 + α → O-16 und O-16 + α → Ne-20 deutlich wird: Bei der ersten werden 7,30 MeV und bei der zweiten nur 4,65 MeV frei; oder wie das Bleiisotop Pb-208 mit $\{N = 126, Z = 82\}$, bei dem die vom Th-232 ausgehende Zerfallsreihe der natürlichen Radioaktivität endet. Auch die anderen natürlichen Zerfallsreihen (↑ 1.1.5.4.) enden bei Bleiisotopen, also bei $Z = 82$.

Die Schaleneffekte sind bei der Gesamtenergie des Kerns, wo sie nur etwa 1 % betragen, nicht allzu deutlich. Einwandfrei nachweisbar sind sie bei den Neutronen- oder Protonen-Separationsenergien, die den Ionisationsenergien der Hülle entsprechen. Bei Schalenabschlüssen treten besonders große, beim Beginn einer neuen Schale besonders kleine Separationsenergien auf. Sehr deutlich spiegeln sich Schaleneffekte in den niederenergetischen *Anregungsspektren* wider, an denen nur wenige „äußere" Nukleonen beteiligt sind. Insbesondere wird die Energie des ersten angeregten Zustands von (g,g)-Kernen extrem groß, wenn N oder Z oder gar beide magische Zahlen sind, z. B. liegt dieser Zustand bei dem doppelt-magischen Bleikern $^{208}_{126}$Pb mit 2,6 MeV etwa sechsmal höher als bei $^{204}_{124}$Hg, da bei der Anregung des Bleikerns voll besetzte Schalen aufgebrochen werden müssen.

Weitere Hinweise ergeben sich aus Häufigkeit und Zahl der stabilen Isotope. So besitzen Zinn $(Z = 50)$ und Blei $(Z = 82)$ einen besonders großen Massenzahlenbereich stabiler Isotope. In beiden Fällen können dem leichtesten stabilen Isotop noch bis zu 12 Neutronen angelagert werden, ohne die Stabilität zu zerstören.

Kerne mit magischen Neutronenzahlen besitzen anomal kleine Einfangquerschnitte für thermische Neutronen, sie zeigen also wenig „Neigung", ihre Neutronenzahl zu verändern.

3.3.3.2. Schalenmodell mit Spin-Bahn-Kopplung. Entsprechend der Nukleonenverteilung im Kern (↑ 3.3.2.1.) ist das auf ein Nukleon wirkende mittlere Kernpotential kugelsymmetrisch, im Kerninneren konstant, der Bindung der Nukleonen entsprechend negativ und im Oberflächenbereich allmählich auf Null ansteigend. Es handelt sich also um einen typischen Potentialtopf (Abb. 3.3.3-1). Die magischen Zahlen sind diejenigen, bei denen jeweils ein Energieniveau bzw. eine Gruppe benachbarter Niveaus entsprechend dem Pauliprinzip gerade voll besetzt ist, bei Beachtung der Entartung aller Zustände im Zentralfeld. Als Ausgangspunkt können die Niveaus des isotropen Oszillators (↑ 2.5.4.6.) dienen; das realistische Kernpotential wird dann innen höher und außen tiefer liegen (↑ Abb. 3.3.3-1). Dementsprechend spalten die Oszillatorniveaus so auf, daß die im Kern innen konzentrierten Zustände mit kleinen Drehimpulsen (↑ 2.5.4.4.) zu größeren Energien und die weiter außen konzentrierten Zustände mit größeren Drehimpulsen zu kleineren Energien verschoben werden. Die so entstehende Niveaufolge (Abb. 3.3.3-2) enthält zwar die schon vom Oszillatorpotential, das tatsächlich für leichte Kerne eine brauchbare Näherung darstellt, gelieferten Schalenabschlüsse bei den ersten drei magischen Zahlen 2, 8 und 20, bietet aber keine Anhaltspunkte für die Erklärung der folgenden. Verfeinerte Betrachtungen bestätigen, daß kein sinnvoller Potentialverlauf die höheren magischen Zahlen liefert.

Die entscheidende Modifikation des Modells besteht in der Einbeziehung einer starken Kopplung zwischen Bahn- und Spindrehimpuls jedes einzelnen Nukleons. Parallel- und Antiparallelstellung beider Drehimpulse, charakterisiert durch die Quantenzahlen $j = l \pm 1/2$ des resultierenden Gesamtdrehimpulses des Nukleons, sind dann energetisch nicht mehr gleichwertig. Die mathematische Beschreibung ist dieselbe wie bei der Erklärung der Feinstruktur (↑ 4.2.2.6.) in der Elektronenhülle, die physika-

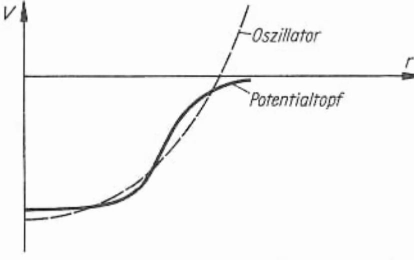


Abb. 3.3.3-1 Oszillatorpotential und realistisches Kernpotential

lische Ursache sind Besonderheiten der Kernkräfte (↑ 3.3.4.2.). Ein noch nicht aufgespaltenes l-Niveau kann wegen der beiden Spineinstellungen $2(2l + 1)$ Protonen oder Neutronen aufnehmen, in dem Niveau mit $j_+ = l + 1/2$ können $2j_+ + 1 = 2(l + 1)$, in dem mit $j_- = l - 1/2$ können $2j_- + 1 = 2l$ untergebracht werden, insgesamt also wieder $2(2l + 1)$. Die Stärke der Aufspaltung wächst mit l bzw. j; das Vorzeichen wird, entgegengesetzt zur Elektronenhülle, so gewählt, daß die Parallelstellung von Bahn- und Spindrehimpuls energetisch günstiger ist (Abb. 3.3.3-3). Dann kann das jeweils tiefste j-Niveau aus dem energetischen Bereich einer Oszillatorschale in den Zwischenbereich, nämlich bei 28, oder sogar in die nächst tiefere Gruppe von Niveaus herabgedrückt werden, nämlich bei den höheren magischen Zahlen 50, 82 und 126 (↑ Abb. 3.3.3-2).

Durch das eben in seinen Grundzügen erläuterte Schalenmodell der Kernstruktur wird ein *nukleares Periodensystem* begründet. Die übliche Bezeichnung *Schalenmodell* deutet hin auf die Analogie zur Elektronenhülle und auf die Existenz von energetischen Abschlüssen, während im Unterschied zur Hülle in der räumlichen Verteilung der Nukleonen keine Schaleneffekte erkennbar sind (↑ 3.3.2.1.).

Voll gefüllte Schalen bzw. Unterschalen, d. h. einzelne j-Niveaus, tragen nichts zum Kernspin I und zum magnetischen Moment bei. Ein Nukleon außerhalb oder ein „Nukleonenloch" in einem sonst vollen Niveau bestimmt den Kernspin allein: $I = j$ mit j als dem Gesamtdrehimpuls des unpaarigen Nukleons oder Lochs. Je zwei Neutronen bzw. je zwei Protonen

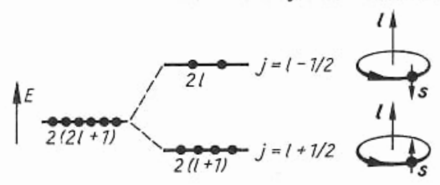


Abb. 3.3.3-3 Zur Aufspaltung durch Spin-Bahn-Kopplung; die Punkte symbolisieren die maximale Nukleonenzahl für $l = 1$

im gleichen j-Niveau stellen im Zustand kleinster Energie ihre Drehimpulse antiparallel, so daß alle (g,g)-Kerne den Grundzustandsspin $I = 0$ haben. Entsprechend wird der Spin aller (u,g)- bzw. (g,u)-Kerne durch das eine unpaarige Nukleon bestimmt. Dies kann zur Festlegung der genauen Niveaufolge herangezogen werden. Viele der niederen Anregungszustände von Kernen mit einem unpaarigen Nukleon können durch alleinige Anregung dieses „Leuchtnukleons" erklärt werden. Aus dem Termschema (↑ Abb. 3.3.3-2) folgt, daß dabei unmittelbar vor Schalenabschlüssen sehr große

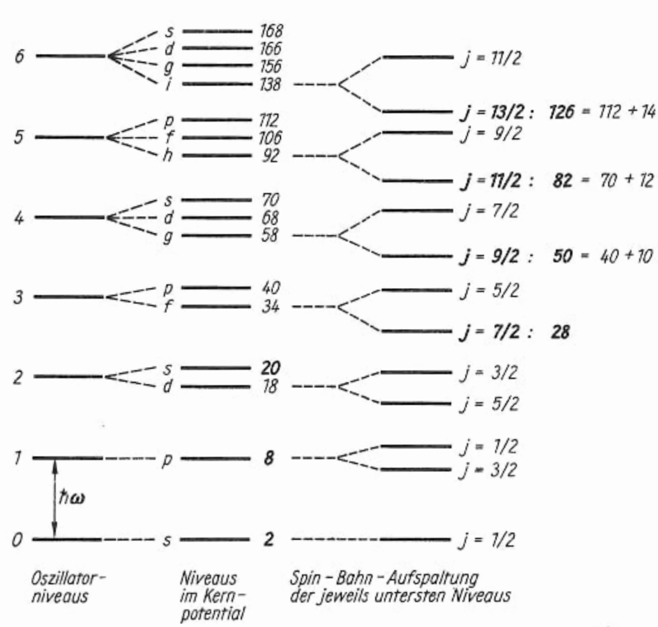


Abb. 3.3.3-2 Schematische Darstellung der Lage der Energieniveaus der Nukleonen; $s, p, d, \dots$ charakterisieren die Bahndrehimpulse $l = 0, 1, 2, \dots$

Spindifferenzen auftreten können. Dies erklärt das gehäufte Auftreten isomerer Kernzustände (↑ 3.3.2.3.2.) im Bereich solcher Nukleonenzahlen, die bisweilen als *Inseln der Isomerie* bezeichnet werden.

Das dargestellte Schalenmodell wird unbrauchbar, wenn sich sehr viele Neutronen und/oder Protonen außerhalb abgeschlossener Schalen befinden. Der aus den abgeschlossenen Schalen gebildete core [*core* engl., Kern] des Kerns wird dann von den Außennukleonen verformt, so daß der gesamte Kern seine Kugelsymmetrie verliert (↑ 3.3.2.1.).

3.3.3.3. Optisches Modell. Gäbe es ausgedehnte Kernmaterie (↑ 3.3.2.2.2.), so könnte man ihre Wirkung auf ein hineingeschossenes Nukleon durch ein mittleres Potential und eine mittlere freie Weglänge (↑ 2.4.3.5.) beschreiben. Der freien Weglänge entspricht im Wellenbild eine räumliche Dämpfung der Welle. Diese Dämpfung wird bewirkt durch die ungeordnete fluktuierende Anordnung der Nukleonen und vor allem durch inelastische Prozesse, bei denen das Geschoßteilchen die Kernmaterie anregt und daher seine Energie verändert. Die Dämpfung bedeutet also nicht, daß das eingeschossene Nukleon verloren geht, sondern nur, daß es den „elastischen Kanal" verläßt, in dem es durch eine Wellenfunktion beschrieben wird. Diese „Absorption" aus dem elastischen Kanal heraus kann auch durch einen imaginären Anteil des Potentials in der Schrödingergleichung oder – wie in der Optik – durch einen komplexen Brechungsindex erfaßt werden.

Die Wechselwirkung eines Nukleons mit einem endlichen Kern ist wegen der diskreten Anregungsniveaus des Compoundkerns (↑ 3.4.2.3.2.) komplizierter als die mit der hypothetischen Kernmaterie. Entsprechend der vorhandenen oder fehlenden Möglichkeit, bei der Anregung des Compoundkerns den Energiesatz zu erfüllen, zeigt der Wirkungsquerschnitt (↑ 2.4.3.5.) und somit auch die Dämpfung der Welle ein ausgeprägtes *Resonanzverhalten* bei Variation der Geschoßenergie (s. a. 3.4.2.1.).

Diese Komplikation verschwindet bei höheren Energien des Nukleons. Bei Prozessen mit Nukleonenenergien von einigen 10 bis einigen 100 MeV wird der Compoundkern so hoch angeregt, daß die einzelnen, aufgrund ihrer endlichen energetischen Breite (↑ 3.3.3.2.) verschwimmenden Niveaus ineinander übergehen, so daß ebenso wie bei einem „Stück Kernmaterie" ein Zustandskontinuum entsteht.

Der relativ leicht meßbare totale Wirkungsquerschnitt für die Wechselwirkung des ankommenden Nukleons mit dem Targetkern hängt nun wesentlich von dem Verhältnis zwischen Kernradius und freier Weglänge ab. Ist, wie früher entsprechend dem *Sandsackmodell* angenommen wurde, die Weglänge klein gegen den Radius, so wirkt der Kern als „schwarzes" Hindernis, das jedes auftreffende Nukleon „absorbiert". Der Wirkungsquerschnitt ist dann von der Größenordnung des geometrischen Querschnitts und zeigt nur eine schwache monotone Energieabhängigkeit. Ist dagegen die Weglänge kleiner oder wenigstens von gleicher Größenordnung wie der Kernradius, so kann die Welle den Kern durchdringen. Dadurch kommt es zu Interferenzeffekten, die zu einer nichtmonotonen Zu- und Abnahme des Wirkungsquerschnitts bei Erhöhung der Geschoßenergie führen. Solche Verläufe werden beobachtet; die Weglänge ist also sogar bei großen Energien nicht klein gegen den Radius.

Mit sinkender Geschoßenergie verringern sich die Möglichkeiten für inelastische Prozesse, so daß die Weglänge wächst. Auch bei niederenergetischen Reaktionen ist daher der Kern durchlässig, er beeinflußt entsprechend dem *optischen Modell* die Welle des Nukleons so wie eine trübe Glaskugel eine Lichtwelle. Die gleichzeitige Anregung der für den jeweiligen Compoundkern charakteristischen Anregungszustände wird mit diesem Bild dadurch vereinbar, daß bei energetischer Mittelung über viele der zugehörigen Resonanzen (Resonanzbreite ≈ 1 eV, Resonanzabstand ≈ 10 eV) die von der Streuung an der trüben Kugel bewirkten *Großresonanzen* sichtbar werden. Diese energetische Mittelung wird automatisch bei Experimenten mit geringer Energieauflösung durchgeführt.

Durch die umfassende Anwendung des optischen Modells zur Interpretation experimenteller Streudaten wurde nachgewiesen, daß sich Nukleonen durch den bzw. in dem Kern relativ frei, d. h. mit einer großen freien Weglänge bewegen können.

3.3.3.4. Kernrotationen. Zahlreiche Daten weisen darauf hin, daß im Bereich der seltenen Erden, für Massenzahlen $150 < A < 190$, und Aktiniden ($A > 222$) die Kerne ihre stabile Kugelform verlieren (↑ 3.3.2.1.3.) und in meist verlängerte Rotationsellipsoide übergehen. Würde ein solcher Kern eine fixierte Orientierung seiner Symmetrieachse haben, so wäre für jedes Nukleon ebenso wie für den gesamten Kern nur noch die Drehimpulskomponente in Achsenrichtung eine Erhaltungsgröße, also ihr quantisierter Wert $\pm \hbar K$ mit K als einer den jeweiligen Kernzustand charakterisierenden inneren Quantenzahl. Für (g,g)-Kerne ist $K = 0$ wegen der paarweisen Antiparallelstellung der Nukleonendrehimpulse (s. a. 3.3.3.2.), für (u,g)- oder (g,u)-Kerne ist K durch das unpaarige Nukleon bestimmt.

Die fehlende Symmetrie des Kerns ermöglicht und erzwingt nun seine Rotation um eine zur Symmetrieachse senkrechte Drehachse, ähnlich wie ein zweiatomiges Molekül um eine Achse

senkrecht zur Molekülachse rotiert (↑ 5.3.1.1.). Der Gesamtdrehimpuls des Kerns besteht aus dem von der „inneren" Bewegung der Nukleonen herrührenden Drehimpuls $\hbar K$ längs der Symmetrieachse und einem mit der kollektiven Rotation des Kerns verbundenen Anteil senkrecht zu dieser Achse. Das Quadrat $\hbar^2 I(I+1)$ des Kerndrehimpulses ist wie üblich quantisiert, wobei $I = K, K+1, K+2, \ldots$ größer oder gleich K ist. Für $K = 0$ führen Symmetrieforderungen ähnlich wie bei zweiatomigen Molekülen dazu, daß nur alle geraden Werte $I = 0, 2, 4, \ldots$ möglich sind.

Mit der Kernrotation ist eine Rotationsenergie verbunden, deren Zusammenhang mit dem Drehimpuls entsprechend

$$(2) \quad E_{\text{rot}} = \frac{I_\perp^2}{2\Theta} = \frac{1}{2\Theta}(I^2 - I_\parallel^2) = \frac{\hbar^2}{2\Theta}[I(I+1) - K^2]$$

durch das *Trägheitsmoment* Θ des Kerns bestimmt ist. $I_\perp$ und $I_\parallel$ sind die Drehimpulskomponenten senkrecht bzw. parallel zu dieser Achse. Auch im Grundzustand, $I = K$, ist $E_{\text{rot}} = \hbar^2 K/2\Theta$ nicht unbedingt Null, was als „Nullpunktsrotation" des Kerns interpretiert werden muß. Durchläuft der Kernspin I die oben angegebenen Werte, so entsteht über der Grundzustandsenergie, evtl. auch über der Energie eines Zustands mit innerer Anregung, eine *Bande von Rotationsniveaus*. Diese zeigen ein einfaches Bildungsgesetz für die Anregungsenergien, nämlich $E_{\text{ang}} = \hbar^2 I(I+1)/2\Theta$, da K für alle Zustände einer Bande denselben Wert hat. Solche kollektiven Anregungen stellen die untersten angeregten Zustände der stark deformierten Kerne dar. Bei Hf-180 verhalten sich die experimentell bestimmten Anregungsenergien wie 1 : 3,3 : 6,9 : 11,7, während die theoretischen Werte 1 : 10/3 : 7 : 12 sind. Bei Ta-181 ist $E_{\text{ang}}^{(11/2)}/E_{\text{ang}}^{(9/2)} = 2{,}21$, der theoretische Wert ist 2,22. Die aus den Anregungsspektren zu entnehmenden Trägheitsmomente sind etwa halb so groß, wie bei starrer Rotation des Kerns um eine Achse senkrecht zur Symmetrieachse zu erwarten wäre.

3.3.4. Kernkräfte

3.3.4.1. Qualitative Merkmale. Als Kernkräfte werden die zwischen den Nukleonen wirkenden Kräfte bezeichnet. Zur Zeit können die Eigenschaften dieser Wechselwirkungen noch nicht ausreichend genau theoretisch berechnet werden. Aus der qualitativ richtigen Vorstellung, daß sie durch den Austausch virtueller, d. h. nicht wirklich als freier Teilchen nachweisbarer π-Mesonen (↑ 3.2.2.2.) zustandekommen, folgt lediglich, daß sie eine *kurze Reichweite* der Ordnung $\hbar/M_\pi c \approx 1$ fm haben, $M_\pi \approx 237\, M_e$ ist die Pionenmasse. Die π-Mesonen sind die leichtesten der an der starken Wechselwirkung (↑ 3.2.4.3.) beteiligten Elementarteilchen. Die angegebene Reichweitenabschätzung folgt daraus, daß ein Pion mit seiner Ruhenergie $M_\pi c^2$ als Mindestenergie trotz der für seine Erzeugung eigentlich gar nicht zur Verfügung stehenden Energie (↑ 2.5.6.3.) für eine Zeitdauer $\Delta t \approx \hbar/M_\pi c^2$ von einem Nukleon emittiert werden kann und daß es in dieser Zeit höchstens eine Strecke $c\Delta t \approx \hbar/M_\pi c$ bis zu dem zweiten Nukleon, wo es wieder absorbiert wird, zurücklegen kann.

Der mittlere Abstand benachbarter Nukleonen im Kern ist durch die Konstante $r_0 \approx 1$ fm in der Kernradienformel (↑ 3.3.2.1.2.) gegeben und stimmt daher größenordnungsmäßig mit der Reichweite der Kernkräfte überein. Ein Nukleon kann also immer nur mit seinen unmittelbaren Nachbarn wechselwirken. Wenn daher, wie das Experiment (↑ 3.3.2.1.2.) zeigt, die Nukleonendichte unabhängig von der Massenzahl ist, so muß dies auch für die Energie pro Nukleon gelten: Das *Sättigungsverhalten* der durch Kernkräfte, bei Vernachlässigung der elektrischen Protonenabstoßung, bewirkten Bindungsenergie pro Nukleon folgt aus der Sättigung der Dichte und der kurzen Reichweite der Kernkräfte.

Die Bindungsenergie beträgt etwa 10 MeV pro Nukleon. Mit der Lokalisierung eines Nukleons im Kernbereich ist aufgrund der Unbestimmtheitsrelation (↑ 2.5.1.7.) eine kinetische Energie der Ordnung 10 MeV verbunden. Die Wechselwirkungspotentiale zwischen den Nukleonen müssen daher eine Tiefe von einigen 10 MeV haben, um die beobachteten Bindungsenergien zu erklären. Diese Wechselwirkungen sind damit natürlich sehr stark, verglichen mit denen in der Atom- oder Molekülphysik. Die Kernwechselwirkungen sind auch stark, verglichen mit der elektrostatischen Abstoßung zwischen den Protonen, da diese andernfalls bei schweren Kernen zu einer verminderten Dichte im Kerninneren und zu einer Anhäufung der Protonen im Bereich der Kernoberfläche führen müßte.

Die Kernwechselwirkungen sind aber schwach, wenn man ihre Energien an den Ruhenergien von Kernteilchen mißt, und zwar nicht nur im Vergleich mit der Ruhenergie $Mc^2 \approx 10^3$ MeV eines Nukleons, sondern auch im Vergleich mit den Ruhenergien $M_\pi c^2 \approx 10^2$ MeV der π-Mesonen. Die auftretenden Energien genügen also nicht zur wirklichen Erzeugung von Teilchen; dies ist der tiefere Grund dafür, daß der Kern nur als Nukleonen aufgebaut gedacht werden kann, zwischen denen in üblicher Weise beschreibbare Kräfte wirken.

Die Kernkräfte erweisen sich auch als schwach

bei der *Bindung zweier Nukleonen* aneinander. Hier existiert nur das aus einem Neutron und einem Proton bestehende *Deuteron* mit der relativ kleinen Bindungsenergie von 2,23 MeV, das keine angeregten Zustände besitzt und in dem die mittlere Entfernung zwischen Proton und Neutron, der Radius des Deuterons, wesentlich größer ist als die Reichweite der anziehenden Wechselwirkung. Dies steht in deutlichem Gegensatz zur Bildung eines Moleküls aus zwei Atomen. Dort entspricht der mittlere Atomabstand ziemlich genau dem Minimum des Anziehungspotentials zwischen den beiden sich fast berührenden Atomen; und die Anziehung ist stark genug, um die Stabilität der Moleküle in einer Vielzahl von Vibrations- und Rotationszuständen zu sichern (↑ 5.3.1.1.).

Aus der für große Massenzahlen A beobachteten Absättigung von Volumen und Bindungsenergie pro Nukleon kann man schließen, daß in der Nukleon-Nukleon-Wechselwirkung abstoßende Elemente vorhanden sein müssen. Anderenfalls wäre trotz des damit verbundenen Anwachsens der kinetischen Energie ein *Kernkollaps* unvermeidbar, bei dem die Kernradien schwerer Kerne unabhängig von der Nukleonenzahl von der Größenordnung der Reichweite der Kernkräfte und die gesamte Bindungsenergie des Kerns $\sim A^2$ wäre.

3.3.4.2. Ergebnisse von Streuexperimenten. Experimentell zugänglich sind die n-p- und p-p-Wechselwirkung über die Untersuchung entsprechender Streuprozesse. Während es unmöglich ist, aus Streuexperimenten mit Energien von einigen MeV auf Einzelheiten der Wechselwirkung zu schließen, konnten durch Experimente bis zu einigen 100 MeV ziemlich genaue Informationen gewonnen werden. Die anfängliche Hoffnung auf ein einfaches Kraftgesetz hat sich dabei allerdings nicht erfüllt.

Die wichtigste allgemeine Erkenntnis ist die *Ladungsunabhängigkeit der Kernkräfte*, d. h., nach Abzug der elektrostatischen p-p-Abstoßung wirken die gleichen Kräfte zwischen p-p, p-n und n-n. Die letzte Aussage ist zwar durch direkte Streuexperimente nicht geprüft, sie folgt aber aus der Gleichwertigkeit von Neutronen und Protonen beim Kernbau, insbesondere aus der Symmetrieenergie in der Energieformel (↑ 3.3.2.2.2.). Die Ladungsunabhängigkeit der Kräfte ist eine starke Stütze für die einheitliche theoretische Behandlung von Neutronen und Protonen als Nukleonen. Aus ihr folgt aber nicht, daß sich die Systeme p-n und n-n, beide ohne Coulombwechselwirkung, gleich verhalten. Ein qualitativer Unterschied zeigt sich bereits in der Existenz des Deuterons als gebundenem n-p-System, und der Nichtexistenz eines Dineutrons oder gar eines Diprotons. Die gemessenen Wirkungsquerschnitte und Winkelverteilungen für p-p- und n-p-Streuung unterscheiden sich ebenfalls wesentlich voneinander, auch wenn der, bei hohen Energien übrigens sehr kleine, elektrostatische Einfluß auf die p-p-Streuung rechnerisch berücksichtigt wird.

Ein Grund für diese zunächst paradox erscheinende Aussage ist die Ununterscheidbarkeit der Protonen, die nach dem Streuprozeß eine Unterscheidung der beiden beteiligten Teilchen verhindert. Darüber hinaus ergeben sich wichtige Einschränkungen aus der für p-p und n-n, aber nicht für n-p zu stellenden Forderung nach einer vollständig antisymmetrischen Wellenfunktion (↑ 2.5.5.4.). Dem Spintriplett-Zustand $S = 1$, anschaulich der Parallelstellung der beiden Nukleonenspins, entspricht eine symmetrische Spinwellenfunktion; dem Singulettzustand $S = 0$, anschaulich der Antiparallelstellung der Spins, eine antisymmetrische Spinwellenfunktion. Die Vertauschung der Nukleonenorte führt zu einem Vorzeichenwechsel ihres Verbindungsvektors $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$. Analysiert man die durch diesen Vektor beschriebene Relativbewegung beider Teilchen nach zugehörigen Werten des Bahndrehimpulses (Quantenzahlen $l = 0, 1, 2, \ldots$), so sind die Zustände mit geradem l-Wert ($l = 0, 2, 4, \ldots$) symmetrisch, die mit ungeradem ($l = 1, 3, 5, \ldots$) dagegen antisymmetrisch gegenüber der Operation $\boldsymbol{r} \to -\boldsymbol{r}$, also gegenüber Vertauschung von $\boldsymbol{r}_1$ und $\boldsymbol{r}_2$. Mit der üblichen Nomenklatur der Zustände (↑ 4.2.4.2.) sind daher entsprechend

$$\text{n-p: } {}^1S,\ {}^3S;\ {}^1P,\ {}^3P;\ {}^1D,\ {}^3D;\ \ldots$$

$$\text{n-n oder p-p: } {}^1S,\ {}^3P,\ {}^1D\ \ldots$$

bei gleichen Nukleonen nur halb so viel Zustände möglich wie im n-p-System. Das Deuteron als das einzige gebundene Zweinukleonsystem ist ein nur für n-p möglicher ${}^3S$-Zustand, allerdings mit einer kleinen Beimengung von ${}^3D$; die Spins von Neutron und Proton stehen also parallel.

Bereits aus der Nichtexistenz des gebundenen Dineutrons folgt, daß die Kräfte im ${}^1S$-Zustand, der als Grundzustand in Betracht käme, weniger als im ${}^3S$-Zustand oder gar nicht anziehend sind. Tatsächlich ist sie nur ein wenig zu schwach, um einen Bindungszustand zu erzeugen. Die *Kernkräfte sind* also *spinabhängig*, sie hängen von der relativen Einstellung der Nukleonenspins, genauer von der Quantenzahl $S = 0$ oder 1 für das Quadrat $\hbar^2 S(S+1)$ des Gesamtspins ab. Diese Aussage wird durch die Streuexperimente bestätigt und präzisiert.

Gleichzeitig zeigen die gemessenen Winkelverteilungen der hochenergetischen n-p-Streuung, daß Vor- und Rückwärtsstreuung (im Schwerpunktssystem) etwa gleich wahrscheinlich sind, obwohl dafür kein zwingender Grund vorliegt. Daraus folgt, daß die Wechselwirkung in P-Zuständen, allgemeiner in solchen mit ungeradem l-Wert, viel schwächer als in S- oder D-Zustän-

den ist. Die Wechselwirkung für ungerade l-Werte verschwindet dann, und zwar unabhängig von den Spineinstellungen, wenn sie proportional zu $1 + \hat{P}^r_{12}$ ist, wobei der auf die Wellenfunktion wirkende Operator $\hat{P}^r_{12}$ die Ortsvektoren beider Teilchen vertauscht. Dies kann anschaulich als Austausch der Teilchen, als Verwandlung des Neutrons in ein Proton beim Stoßprozeß interpretiert werden. Solche *Austauschkräfte* wurden bereits anfangs der 30er Jahre von HEISENBERG und MAJORANA vorgeschlagen; sie können über den Austausch geladener π-Mesonen auch physikalisch verstanden werden.

Der experimentell gefundene Anteil von etwa 50% Austauschkräften genügt nicht, um abstoßende Wechselwirkungselemente und damit eine Verhinderung des Kernkollaps (↑ 3.3.4.1.) zu bewirken. Dagegen konnte die Existenz eines sogenannten hard core [*hard core* engl., harter Kern] wahrscheinlich gemacht werden. Darunter ist ein Bereich sehr starker Abstoßung bei kleinen Nukleonenabständen zu verstehen. Die Nukleonen verhalten sich wie harte undurchdringliche Kugeln, wenn ein Abstand von etwa 0,5 fm unterschritten wird, d. h., jedes als harte Kugel idealisierte Nukleon besitzt einen *Nukleonenradius* von 0,2 bis 0,3 fm. Dieses *hard-core-Modell* der Kernkräfte garantiert die Sättigung der Nukleonendichte bei großen A. Bemerkenswerterweise ist aber die Kernmaterie außerordentlich locker gepackt. Während in einem Molekül so wie in einer Flüssigkeit oder in einem Kristall die einzelnen Atome bzw. Ionen sich praktisch berühren, ist das Verhältnis des oben definierten Nukleonenradius zum mittleren Abstand etwa $0{,}25/1{,}2 \approx 1/5$, das von den hard cores erfüllte Volumen also nur etwa 1% des zur Verfügung stehenden Kernvolumens. Die eigenartige „Schwäche" der anziehenden Kernwechselwirkungen zeigte sich bereits beim Deuteron (↑ 3.3.4.1.). Die Kernstruktur ist gasartig, womit ein neues Licht auf den Erfolg des Modells unabhängiger Nukleonen (↑ 3.3.3.1. bis 3.3.3.3.) fällt.

Mit Spinabhängigkeit, Austauschanteil und hard core sind die Eigenschaften der Kernkräfte noch nicht erschöpft. Aus den Streudaten und aus dem endlichen elektrischen Quadrupolmoment des Deuterons folgt, daß die Kernkräfte nichtzentrale Anteile enthalten: Die Kräfte wirken nicht nur in Richtung der Verbindungslinie beider Nukleonen. Solche Wechselwirkungen sind immer möglich, wenn die miteinander wechselwirkenden Teilchen selbst eine Richtung auszeichnen, so wie es hier durch den Nukleonenspin gegeben ist. Die Wechselwirkungsenergie hängt dann von der Orientierung der Spins bzw. Momente in bezug auf die Verbindungslinie beider Teilchen ab.

Außerdem zeigen die Experimente die Existenz einer *Spin-Bahn-Kopplung*, d. h. eine Abhängigkeit der Wechselwirkung von der relativen Orientierung zwischen Gesamtbahn- und Gesamtspindrehimpuls der beiden Nukleonen. Diese Anteile der Kernkräfte tragen ebenso wie die nichtzentralen zu der im Schalenmodell empirisch eingeführten Spin-Bahn-Kopplung bei (↑ 3.3.3.2.).

3.4. Kernumwandlungen

Bereits vor Erkenntnissen und Modellen über Struktur und Verhalten der Kerne war durch die Untersuchung der natürlichen und die Erzeugung der künstlichen Radioaktivität (↑ 1.1.5.5.) bekannt, daß sich die Kerne ineinander umwandeln können. Während zur Einleitung von Kernumwandlungen zunächst nur die natürliche Radioaktivität zur Verfügung stand, ergaben sich mit der Entwicklung der Teilchenbeschleuniger (↑ 3.1.) für geladene Teilchen und der Kernreaktoren (↑ 3.4.2). als intensive Neutronenquellen vielfältige neue Möglichkeiten. Fast immer entstehen als Zwischenprodukte instabile Kerne, deren Zerfälle untersucht werden.

Der historisch entstandene Ausdruck „Kernzerfall" ist eigentlich nur bei der Spaltung schwerer und beim α-Zerfall einiger leichter Kerne, etwa von Be-8, zutreffend. Sonst ändert sich die Zahl der Nukleonen im Kern entweder gar nicht (β-Prozesse) oder relativ wenig. Die Abgrenzung zwischen den Reaktionen bei zwei aufeinanderprallenden Kernen und dem „Zerfall" eines einzelnen Kerns ist fließend, da eigentlich jeder instabile Kern nur ein Zwischenprodukt einer – unter Umständen vor sehr langer Zeit eingeleiteten – Kernreaktion ist. Im folgenden werden, in Übereinstimmung mit dem allgemeinen Sprachgebrauch, nur die Umwandlungen von nicht hochangeregten Kernen als Kernzerfälle bezeichnet.

3.4.1. Kernzerfälle

Die Erscheinungen der α-, β -und γ-Umwandlungen wurden bereits bei der natürlichen Radioaktivität (↑ 1.1.5.1.) gefunden. Zusammen mit der Kernspaltung sind sie für das Verständnis der Stabilität der Kerne (↑ 3.3.1.2.) und die Aufklärung ihrer Termschemata (↑ 3.3.2.3.1.) wesentlich. Im folgenden werden diese Prozesse genauer charakterisiert.

3.4.1.1. Elektromagnetische Prozesse

3.4.1.1.1. *Emission von γ-Quanten.* Alle allgemeinen Aussagen über die Natur der Strahlungsprozesse (↑ 4.1.) bleiben bei Übergängen zwischen den verschiedenen diskreten Energieniveaus eines Kerns unverändert gültig. Die

emittierte elektromagnetische Strahlung wird γ-Strahlung, die zugehörigen energiereichen Photonen werden γ-Quanten genannt. Bei Photonenenergien im MeV-Bereich sind die zugehörigen Wellenlängen so klein, daß sie nicht mehr direkt gemessen werden können. Daher ist die kernphysikalische Gammaspektroskopie auf Messungen der Photonenenergien angewiesen. Die energieärmeren γ-Quanten, etwa mit Energien im 100-keV-Bereich, entsprechen völlig den Quanten entsprechend kurzwelliger Röntgenstrahlung.

Die möglichen *Strahlungstypen* können nach sogenannten Multipolordnungen $l = 1, 2, 3, \ldots$ eingeteilt werden, die als Dipol-, Quadrupol-, Oktupolstrahlung usw. bezeichnet werden. l kann als Drehimpulsquantenzahl (↑ 2.5.3.2.) des Drehimpulses interpretiert werden, den das Photon mit fortträgt. Es gibt keine Photonen ohne Drehimpuls, so daß es keine zu $l = 0$ gehörige Monopolstrahlung gibt. Mit wachsender Multipolordnung wird die Winkelverteilung der emittierten Strahlung komplizierter, während die Übergangswahrscheinlichkeiten rasch abnehmen, und zwar größenordnungsmäßig um jeweils einen Faktor $(R/\lambda)^2$ mit $R \approx 10^{-14}$ bis 10^{-15} m als dem Kernradius und $\lambda \approx 10^{-12}$ bis 10^{-13} m als typischen Wellenlängen. Charakteristische Werte für das Verhältnis der Übergangswahrscheinlichkeiten benachbarter Multipolordnungen sind 10^{-3} (Quantenenergie ≈ 1 MeV) bis 10^{-5} (Quantenenergie ≈ 100 keV). Dennoch treten höhere Multipolstrahlungen als Folge der Drehimpulserhaltung beim Strahlungsprozeß auf. Sind I_a und I_b die Kernspins der beiden miteinander kombinierenden Kernniveaus, so muß nach den quantenmechanischen Regeln für die Vektoraddition von Drehimpulsen (↑ 2.5.3.4.) gemäß $|I_a - I_b| \leqq l \leqq I_a + I_b$ die Multipolordnung mindestens gleich der Differenz der Kernspins sein. Daher können entsprechend unwahrscheinliche und daher intensitätsschwache Übergänge die einzige Möglichkeit zur Energieabgabe aus einem langlebigen Kernniveau sein. Solche Niveaus und mit ihnen auch die emittierte γ-Strahlung haben eine außergewöhnlich kleine Energieunschärfe (↑ 3.3.2.3.). Darauf beruht die extreme Meßgenauigkeit beim Mößbauereffekt (8.1.2.6.).

Strahlungsübergänge höherer Multipolordnung treten prinzipiell auch in der Atomhülle auf. Sie sind jedoch im allgemeinen nicht beobachtbar und werden daher als verbotene Übergänge klassifiziert, weil die angeregten Atome ihre Energie bereits vorher bei Zusammenstößen mit anderen Atomen abgeben. Dagegen fehlen entsprechende störende Einflüsse bei den Kernen, weshalb hier auch sehr unwahrscheinliche Übergänge nachweisbar sind.

Innerhalb jeder Ordnung l wird noch zwischen elektrischer und magnetischer Multipolstrahlung unterschieden. Beispielsweise kann die mit $E1$ symbolisierte elektrische Dipolstrahlung, die in der Elektronenhülle praktisch allein auftritt, im klassischen Bild auf einen schwingenden elektrischen, entsprechend die magnetische Dipolstrahlung $M1$ auf einen schwingenden magnetischen Dipol zurückgeführt werden. Wichtig ist die elektrische Quadrupolstrahlung $E2$, die für die Übergänge zwischen benachbarten Rotationsniveaus deformierter Kerne (↑ 3.3.3.4.) verantwortlich ist. Dieser Strahlungstyp würde auch klassisch von einem rotierenden geladenen Rotationsellipsoid ausgesandt werden.

3.4.1.1.2. *Innere Konversion.* Die Anregungsenergie des Kerns kann auch direkt auf ein Hüllenelektron übertragen werden, das dadurch mit hoher Energie aus dem Atomverband herausgeschleudert wird. Da dafür verschieden fest gebundene Elektronen aus den inneren Schalen der Atomhülle in Betracht kommen, besitzen die einem Kernübergang zuzuordnenden *Konversionselektronen* etwas unterschiedliche Energien. Die Bezeichnung innere Konversion oder innere Umkehrung ist irreführend; sie entstand aus der Vorstellung, daß der Kern zunächst ein γ-Quant emittiert, das dann wieder von einem Hüllenelektron absorbiert wird. Tatsächlich handelt es sich aber um einen auch „im Inneren" strahlungslosen Prozeß, dessen Übergangswahrscheinlichkeit proportional ist zur Wahrscheinlichkeit, mit der sich ein Hüllenelektron im Kern aufhält. Diese Wahrscheinlichkeit ist bei schweren Kernen wegen der mit der großen Kernladung verbundenen starken Kern-Elektronen-Anziehung besonders groß. Die innere Konversion als Konkurrenzprozeß zur Emission eines γ-Quants ist bei schweren Kernen, kleinen Übergangsenergien und hohen Multipolordnungen effektiv. Der *Konversionskoeffizient*, der als Verhältnis der Übergangswahrscheinlichkeiten für Konversion und Strahlung definiert ist, hängt nicht von Details der Kernstruktur ab. Seine Messung kann daher zur Bestimmung des Strahlungstyps und damit zur Festlegung von Kernspins und -paritäten oder auch zur Ermittlung der Kernladungszahl dienen.

Die bei der Konversion in einer inneren Elektronenschale entstehende Lücke führt beim „Nachspringen" äußerer Elektronen in diese Lücke zur Emission charakteristischer Röntgenstrahlung des betreffenden Elements.

Bei leichten Kernen und großen Übergangsenergien kann auch der seltene Prozeß beobachtet werden, bei dem die frei werdende Energie zur Erzeugung eines Elektron-Positron-Paars dient, wozu mindestens die doppelte Ruhenergie $2\,M_e c^2 \approx 1$ MeV eines Elektrons notwendig ist. Auch dieser Prozeß der sogenannten *inneren Paarbildung* erfolgt ohne Vermittlung eines γ-Quants.

3.4.1.2. Betazerfall

3.4.1.2.1. *Einführung des Neutrinos.* Die Emission der sogenannten β-Strahlung wurde bereits bei den frühen Versuchen mit radioaktiven Präparaten festgestellt. Die ausgesandten leichten und negativ geladenen Teilchen wurden bald als Elektronen identifiziert. Bei einem Kernprozeß wird ein Elektron emittiert, das aber ebenso wenig wie ein γ-Quant bereits vorher im Kern vorhanden war. Die Erhaltung der elektrischen Ladung wird dadurch gesichert, daß sich ein Neutron in ein Proton umwandelt. Da die Neutronenmasse um etwa 2,5 Elektronenmassen über der des Protons liegt, ist das freie, nicht im Kern gebundene Neutron instabil (↑ 3.4.4.1.).

Im Unterschied zur diskreten, für einen gegebenen Kernübergang genau bestimmten Energie eines γ-Quants oder eines Konversionselektrons (↑ 3.4.1.1.2.) haben die beim β-Zerfall emittierten Elektronen Energien, die kontinuierlich zwischen Null und einer für den betrachteten β-Übergang charakteristischen Maximalenergie variieren. Kann der β-Übergang sowohl zum Grundzustand als auch zu einem angeregten Zustand des isobaren Nachbarkerns verlaufen, und geht dieser angeregte Zustand durch Emission eines γ-Quants der Energie E_γ in den Grundzustand über, so hängen die Maximalenergien der beiden β-Prozesse über $E_{\max}^{\text{grund}} = E_{\max}^{\text{ang}} + E_\gamma$ miteinander zusammen. Dies zeigt, daß der *Energiesatz* auch für β-Prozesse gilt, und daß die Maximalenergie des Elektrons die beim jeweiligen Übergang frei werdende Energie darstellt. Wenn dennoch das Elektron fast immer weniger Energie erhält, so muß ein weiteres emittiertes Teilchen die fehlende Energie erhalten haben. Dieses Teilchen ist das 1934 von PAULI postulierte *Neutrino.* Es ist ungeladen und hat keine Ruhmasse, ähnlich dem Photon. Im Unterschied zu diesem hat es den Spin $s = 1/2$.

Nicht nur der Energie-, sondern auch der *Drehimpulssatz* werden durch die Einführung des Neutrinos „gerettet". Denn die Emission eines Elektrons als eines Teilchens mit dem Spin $s = 1/2$ wäre nur mit einer Spindifferenz von $\Delta I = 1/2, 3/2, \ldots$ zwischen Anfangs- und Endkern vereinbar, während doch alle isobaren Kerne in allen ihren Zuständen entweder ganzzahlige (A gerade) oder halbzahlige (A ungerade) Spinwerte haben (↑ 3.3.2.4.). Das Problem ist gelöst, wenn das Neutrino ebenso wie das Elektron oder die Nukleonen einen halbzahligen Spin hat.

Neutrinos sind sehr schwer nachweisbar, da sie riesige Materieschichten praktisch wechselwirkungsfrei durchdringen. Indirekte Nachweise gelangen durch die Messung des Rückstoßes des sich umwandelnden Kerns, also durch Anwendung des Impulssatzes. Der direkte Nachweis von (Anti-) Neutrinos erfolgte 1956 (↑ 3.2.2.1.).

3.4.1.2.2. *Grundprozesse.* Die Grundprozesse des β-Zerfalls sind die Umwandlung der Nukleonen nach den zwei Reaktionsgleichungen $n \to p + e^- + \bar{\nu}$ und $p \to n + e^+ + \nu$. Dabei bezeichnen e^- das Elektron und e^+ das Positron als Antiteilchen (↑ 3.2.1.7.) des Elektrons, ν und $\bar{\nu}$ Neutrino und Antineutrino. Ebenfalls möglich sind nach einem allgemeinen Prinzip der Elementarteilchenphysik (↑ 3.2.1.7.), wonach stets Teilchen auf einer Seite einer Reaktionsgleichung durch ihre Antiteilchen auf der anderen Seite ersetzt werden können, die Prozesse $n + e^+ \to p + \bar{\nu}$ und $p + e^- \to n + \nu$. Die *Positronenemission* ist zwar durch ein freies Proton wegen der größeren Masse und damit auch größeren Ruhenergie des Neutrons unmöglich, kann aber bei Einbettung des Protons in einen Kern energetisch möglich werden (↑ 3.3.1.2. und 3.3.2.2.4.). Der Positroneneinfang scheidet normalerweise aus, weil in der Umgebung des Kerns keine Positronen vorhanden sind. Der *Elektroneneinfang* ist dagegen möglich. Er ist um die Summe der Ruhenergien von Elektron und Positron, wegen der Massengleichheit beider Teilchen also um die doppelte Ruhenergie des Elektrons (etwa 1 MeV), günstiger als der Konkurrenzprozeß „Positronenemssion". Bei schweren Kernen stehen die kernnahen Hüllenelektronen mit ausreichender Wahrscheinlichkeit im Kernbereich zum Einfang zur Verfügung. Daher dominieren bei solchen Kernen die Elektroneneinfangprozesse. Am wahrscheinlichsten ist der als K-Einfang bezeichnete Einfang aus der innersten, der K-Schale. Ebenso wie bei der inneren Konversion (↑ 3.4.1.1.2.) tritt Röntgenstrahlung als Folgeprozeß auf.

β^--Zerfall tritt bei Kernen mit zuviel Neutronen auf. Solche Kerne entstehen bei Neutroneneinfang und Prozessen der Massenzahlverkleinerung, also bei α-Zerfall und Kernspaltung, da schwerere Kerne einen größeren relativen Neutronenüberschuß als leichtere Kerne haben (↑ 3.3.1.3.). Besonders bei der Spaltung entstehen primär Tochterkerne weitab von der Stabilitätslinie, die allerdings so hoch angeregt sind, daß sie ihren Neutronenüberschuß größtenteils direkt durch den schnellen Prozeß der Neutronenemission abbauen können. Dennoch treten auch hier noch β^--Zerfälle auf, nachdem die Kerne durch Neutronenverdampfung hinreichend „abgekühlt" sind. Bei dem gegenüber der Spaltung weniger drastischen Vorgang des α-Zerfalls aus nicht oder schwach angeregten Mutterkernen treten hoch angeregte Endkerne gar nicht erst auf, so daß die Elektronenemission der einzige Prozeß ist, der die anschließenden isobaren Umwandlungen zur Herausbildung stabiler Kerne ermöglicht.

Kerne mit zu viel Protonen kommen in der Natur unter irdischen Bedingungen nicht vor. Sie entstehen immer dann, wenn schwerere Kerne aufgebaut werden, also bei der Beschießung von Kernen mit möglichst massereichen Teilchen. Eine wichtige Rolle haben β^+-Prozesse und Elektroneneinfang beim Aufbau der mittleren und schweren Elemente in einer frühen Phase der Entwicklung des Universums gespielt.

3.4.1.2.3. *Energieverteilung und Zerfallswahrscheinlichkeiten.* Ebenso wie bei γ-Übergängen (↑ 3.4.1.1.1.) nimmt auch beim β-Zerfall die Zerfallswahrscheinlichkeit mit wachsender Spindifferenz ΔI der kombinierenden Niveaus rasch ab. Als *erlaubte Übergänge* werden diejenigen

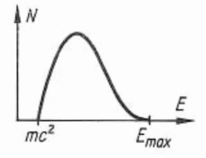


Abb. 3.4.1-1 Energieverteilung der Elektronen bzw. Positronen beim β-Zerfall; Erläuterung im Text

bezeichnet, bei denen $\Delta I = 0$ oder 1 ist; sie können durch Elektronen- und Neutrinospin allein, also ohne Beteiligung von Bahndrehimpulsen bilanziert werden. Das beobachtete Energiespektrum der emittierten Elektronen bzw. Positronen ergibt sich bei solchen Übergängen aus der Annahme, daß alle mit der Energiebilanz verträglichen Endzustände, bei denen also Elektron (Positron) und Antineutrino (Neutrino) mit unterschiedlichen Impulsen in verschiedene Richtungen emittiert werden, gleich wahrscheinlich sind. Die Zahl $N(E)$ der Elektronen (Positronen) pro Energieintervall ergibt sich zu $N(E) \sim E\sqrt{E^2-(M_ec^2)^2}\,(E_{max}-E)^2$, $E \geqq M_ec^2$ ist die Elektronenenergie einschließlich der Ruhenergie, E_{max} ihr Maximalwert. Der nicht angegebene Proportionalitätsfaktor ist für alle erlaubten Übergänge von gleicher Größenordnung. Die Form der Funktion $N(E)$ (Abb. 3.4.1-1) stimmt gut mit den gemessenen Betaspektren überein. Das Integral $f(E_{max}) = \int_{M_ec^2}^{E_{max}} \mathrm{d}E N(E)$ über $N(E)$, also die Gesamtzahl der emittierten Elektronen, ist ein Maß für die Übergangswahrscheinlichkeit γ des Prozesses. Tatsächlich erweisen sich die als $f\tau$-Werte bezeichneten Produkte aus dem von der Maximalenergie abhängenden Wert der Funktion f und der mittleren Lebensdauer $\tau = \gamma^{-1}$ als etwa gleich für alle erlaubten Übergänge, obwohl τ selbst um viele Zehnerpotenzen variieren kann. Beispielsweise hat der $f\tau$-Wert für den mit einer mittleren Lebensdauer von etwa 10^3 s erfolgenden Zerfall des freien Neutrons die gleiche Größenordnung wie für den β-Zerfall des Tritons H-3 mit $\tau \approx 10^8$ s.

Die Energieverteilung $N(E)$ gilt für die den Kern gerade verlassenden Elektronen bzw. Positronen. Danach werden die Elektronen durch die vom entstandenen Endkern ausgehende elektrische Anziehung verzögert, während die Positronen beschleunigt werden. Dadurch verändert sich die Energieverteilung vor allem im niederenergetischen Bereich.

3.4.1.2.4. *Verlust der Spiegelsymmetrie beim β-Zerfall.* 1956 formulierten Lee und Yang im Zusammenhang mit Problemen der Elementarteilchenphysik die Hypothese der *Nichterhaltung der Parität*, also der Verletzung der räumlichen Spiegelsymmetrie bei Prozessen der schwachen Wechselwirkung. Vorher galt als selbstverständlich, daß zu jedem physikalischen Prozeß auch der räumlich gespiegelte Prozeß möglich sei. Bei einer solchen Spiegelung gehen Rechts- in Linksschrauben über (s. a. 2.1.2.7.). Bereits im gleichen Jahr 1956 wurde von Wu und Mitarbeitern in einem berühmten Experiment die *Auszeichnung eines bestimmten Schraubensinns* beim β-Zerfall demonstriert. Dazu wurden die Richtungen registriert, unter denen die Elektronen von einem radioaktiven Co-60-Präparat beim Zerfall Co-60 → Ni-60 + e^- + $\bar{\nu}$ ausgesandt wurden, wobei die magnetischen Momente und damit auch die Spins der Kobaltkerne durch ein starkes Magnetfeld bei tiefen Temperaturen polarisiert, also bevorzugt parallel zum Magnetfeld orientiert wurden. Das Experiment zeigt, daß die Elektronen bevorzugt in die entgegengesetzte Richtung emittiert werden. Da der Kernspin wie jeder Drehimpuls ein axialer Vektor (↑ 2.1.2.7.) ist, bedeutet dies, daß Drehsinn der Kobaltkerne und bevorzugte Emissionsrichtung der Elektronen zusammen eine Linksschraube bilden. Dies gilt auch für andere β-Prozesse, während bei Positronenemission eine Rechtsschraubenzuordnung auftritt.

3.4.1.3. Alphazerfall und spontane Kernspaltung

3.4.1.3.1. *Metastabilität und Coulombwall.* Dem α-Zerfall und der Spaltung gemeinsam ist die spontane, d. h. ohne äußere Anregung vor sich gehende Neuaufteilung der Nukleonen des Mutterkerns auf zwei Tochterkerne. Bei der

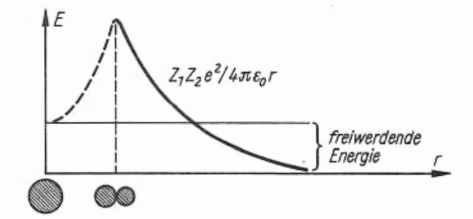


Abb. 3.4.1-2 Zum Coulombwall bei α-Zerfall oder Spaltung für einen energetisch instabilen Kern (Erläuterung im Text)

Spaltung haben diese vergleichbare Massen, wobei das Massenverhältnis in ziemlich weiten Grenzen variieren kann und die Spaltung in zwei gleich große Bruchstücke nicht den wahrscheinlichsten Prozeß darstellen muß (Abb. 3.4.1-2). Beim α-Zerfall eines schweren Kerns sind die beiden Bruchstücke sehr verschieden groß: Das kleine ist ein α-Teilchen, also ein aus 2 Neutronen und 2 Protonen bestehender He-4-Kern. Daher kann man sagen, daß der Ausgangskern das α-Teilchen bei seinem Zerfall „emittiert", es ist aber in diesem vor dem Zerfall so wenig als identifizierbare Untereinheit enthalten wie die beiden Endkerne in dem sich spaltenden Ausgangskern.

Die *energetische Instabilität* schwerer Kerne beruht auf der durch die Coulombsche Abstoßung zwischen den Protonen bewirkten Abnahme der Bindungsenergie pro Nukleon mit wachsender Massenzahl (↑ 3.3.2.2.1.). Bereits bei mittleren Massenzahlen um $A \approx 150$ würde beim α-Zerfall ebenso wie bei der Spaltung Energie frei. Beim α-Zerfall ist die überwiegende Mehrzahl der Nukleonen im leichteren Endkern etwas fester gebunden und die vier Nukleonen des α-Teilchens mit der für einen so leichten Kern sehr großen Bindungsenergie von reichlich 7 MeV je Nukleon sind nur wenig lockerer gebunden als im Ausgangskern. Bei der symmetrischen Spaltung von U-238 werden etwa 180 MeV, bei der Umwandlung der primären Spaltprodukte in stabile Endkerne nochmals etwa 30 MeV frei. Aber selbst für Uran ist die Spaltwahrscheinlichkeit noch wesentlich kleiner als die ohnehin schon kleine für α-Zerfall, so daß die *spontane Spaltung* 1940 von FLEROW erst *nach* der neutroneninduzierten Spaltung entdeckt wurde.

Die Metastabilität mittelschwerer Kerne beruht darauf, daß beim Zerfallsprozeß energetisch ungünstige Zwischenstadien auftreten. Dies wird besonders deutlich, wenn man den umgekehrten Prozeß der Annäherung und schließlichen Verschmelzung zweier Teilkerne mit den Protonenzahlen Z_1 und Z_2 betrachtet. Diese stoßen sich gegenseitig ab, so daß bei ihrer Annäherung ständig Arbeit verrichtet werden muß. Sie besitzen eine potentielle Energie $E_p = Z_1Z_2e^2/4\pi\varepsilon_0 r$ (↑ 2.1.5.5.), die mit kleiner werdendem Abstand r ihrer Mittelpunkte anwächst. Sie erreicht ihren Maximalwert $Z_1Z_2e^2/4\pi\varepsilon_0(R_1 + R_2)$, wenn sich die beiden Teilkerne mit den Radien R_1 und R_2 berühren (↑ Abb. 3.4.1-2). Etwa von da an wirken die anziehenden kurzreichweitigen Kernkräfte, die Teile verschmelzen. Diese Prozeßstadien können nicht mehr durch einen einfachen Parameter wie den Abstand r charakterisiert werden, aber am Schluß liegt der kugelsymmetrische Verbundkern vor. Seine Energie ist bei der gewählten Energieskala, in der den getrennten, ruhenden Teilkernen die Energie Null zugeordnet wurde, positiv oder negativ, je nachdem, ob bei seiner Auftrennung Energie frei wird oder verbraucht wird, bei den gegenüber α-Zerfall bzw. Spaltung metastabilen Kernen also positiv.

Wäre der Kern ein klassisches System, so hätte er keinerlei Möglichkeit, von sich aus, d. h. ohne Zufuhr der zur Überwindung der Coulombbarriere notwendigen Aktivierungsenergie, zu zerfallen. Wird diese Energie zugeführt, so zerfällt der Kern praktisch sofort. Dies geschieht im Kernreaktor bei der induzierten Kernspaltung. Die spontane Spaltung beruht dagegen ebenso wie der α-Zerfall auf dem quantenmechanischen Tunneleffekt (↑ 2.5.4.8.), der es nicht nur einem einzelnen Teilchen, sondern auch komplizierteren quantenmechanischen Systemen erlaubt, klassisch verbotene Zwischenbereiche zu überwinden, gewissermaßen die Coulombbarriere zu durchtunneln. Die Wahrscheinlichkeit für diesen Tunnelprozeß bestimmt entscheidend die Größenordnung der Zufallswahrscheinlichkeit. Sie hängt bei gegebener Barrierenform und -höhe außerordentlich stark von der Zerfallsenergie ab. Daher werden emittierte α-Teilchen nur in dem kleinen Energiebereich zwischen 4 und 9 MeV beobachtet. Die langlebigen Uran- und Thoriumkerne, deren mittlere Lebensdauern mit dem Alter der Erde vergleichbar sind, haben α-Energien knapp über 4 MeV, während das Isotop Rn-215 des radioaktiven Edelgases Radon mit $E = 8{,}8$ MeV bereits in 10^{-6} s zerfällt. Die extreme Energieabhängigkeit der Tunnelwahrscheinlichkeit erklärt auch, daß der α-Zerfall fast immer zum Grundzustand des Endkerns und nur selten zu dessen angeregten Zuständen führt.

3.4.1.3.2. *Stabilität gegen Spaltung.* Die Höhe der Barriere ist bei der Spaltung eines schweren Kerns in zwei etwa gleich große Bruchstücke wesentlich niedriger, als es der Abschätzung mittels der sich berührenden Teilkerne entspricht. Die Spaltung verläuft in Wirklichkeit so, daß sich der Ausgangskern zunächst verlängert, dann bei gleichzeitiger weiterer Vergrößerung der Gesamtlänge mehr und mehr einschnürt und schließlich in zwei zunächst nicht kugelförmige Teile zerreißt. Die *kritische Form* bestimmt die Höhe der Barriere bzw. die Größe der Aktivierungsenergie. Diese liegt für Uran- und Plutoniumkerne im Bereich 5 bis 6 MeV, gegenüber 70 bis 80 MeV bei zwei sich berührenden Kugeln. Nur dadurch wird die Spaltung überhaupt konkurrenzfähig mit dem α-Zerfall.

Der Ausgangskern wird stabilisiert durch die mit jeder Deformation verbundene Vergrößerung der zu $A^{2/3}$ proportionalen Oberflächenenergie (↑ 3.3.2.2.); als treibende, den Kern destabilisierende Kraft der Spaltprozesse wirkt die elektrostatische Abstoßung. Dabei ist das Verhältnis der zu $Z^2/A^{1/3}$ proportionalen elektrostatischen Energie zur Oberflächenenergie, also

der Parameter Z^2/A, die entscheidende Größe für die Spaltbarkeit. Eine energetische Spaltbarriere bildet sich nur, solange der Kern wenigstens stabil gegen kleine Deformationen ist, d. h., daß bei einer kleinen Verformung der Ausgangskugel die Oberflächenenergie stärker anwächst als die elektrostatische Energie abnimmt. Für $Z^2/A \geq 50$ wird der Kern bereits gegen kleine Deformationen instabil, d. h., er zerfällt sofort. Nach dieser Abschätzung können Kerne, deren Massenzahlen (bei Berücksichtigung des Neutronenüberschusses) wesentlich über $A \approx 300$ liegen, nicht existieren. Die durch spontane Spaltung bedingten rasch abnehmenden Lebensdauern der künstlich hergestellten Transurane bestätigen diese Schlußfolgerung.

3.4.2. Kernreaktionen

3.4.2.1. Allgemeiner Ablauf und Klassifizierung. Die Wechselwirkung von Atomkernen untereinander geschieht über die kurzreichweitigen, aber starken Kernkräfte und die langreichweitigen elektromagnetischen Kräfte. Aufgrund der relativ scharfen Begrenzung eines Atomkerns wirkt in größeren Abständen zunächst nur die elektrostatische Coulombkraft. Sie führt zur elastischen Streuung der stoßenden Teilchen (Rutherfordstreuung) sowie zur Energieübertragung, welche die Anregung eines oder beider Stoßpartner zur Folge hat (Coulombanregung).

Eigentliche Kernreaktionen mit Änderung der Nukleonenzahl der Stoßpartner treten erst beim gegenseitigen Eindringen der Stoßpartner in den Bereich ihrer Kernkräfte, d. h. bei Annäherung auf einige Fermi (1 fm = 10^{-15} m), ein. Geladene Geschoßteilchen benötigen deshalb wenigstens die kinetische Energie zur Überwindung der Coulombschwelle (↑ 3.4.1.3.1), während die ungeladenen Neutronen bereits mit geringsten kinetischen Energien Kernreaktionen auslösen.

Entsprechend den Gesetzen der Quantenmechanik ist der klassische Bahnbegriff für die Reaktionspartner im allgemeinen nicht anwendbar. So führt der Tunneleffekt bereits unterhalb der Coulombschwelle zur Auslösung von Kernreaktionen.

Symbolisch werden Kernreaktionen wie in einer chemischen Reaktionsgleichung durch die reagierenden Teilchen (A, a), die Produkte (B, b) sowie die sogenannte Wärmetönung Q der Reaktion, d. h. die freiwerdende Energie, beschrieben. Es gilt $A + a \to B + b + Q$; A wird als Targetkern bezeichnet. Die Kurzbezeichnung obiger Reaktion ist $A(a, b)B$. Beim Stoß zweier Kerne, wozu auch Protonen und Neutronen zählen, können mehrere Reaktionstypen auftreten: 1. die elastische Streuung $A(a, a)A$, 2. die inelastische Streuung $A(a, a')A^*$ (A^* angeregter Targetkern), 3. der Strahlungseinfang $A(a, \gamma)B$ (γ: emittierter γ-Quant), 4. ein Teilchentransfer $A(a, b)B$ (Austausch bzw. Übergang von Nukleonen) und 5. Mehrteilchenreaktionen $A(a, bc)B$.

Bei Kernreaktionen bleiben Ladungs- und Nukleonenzahl (letztere im konventionellen Energiebereich bis zu Energien von etwa 100 MeV/u), Gesamtdrehimpuls und Parität erhalten. Die Wärmetönung ergibt sich entsprechend der Energieerhaltung aus den Massen der beteiligten Kerne. Für die Reaktion $A(a, b)B$ gilt $Q = E_B(B) + E_B(b) - E_B(A) - E_B(a)$ mit den Bindungsenergien E_B der Reaktionspartner. Der Q-Wert tritt bei exothermen Reaktionen ($Q > 0$) als kinetische Energie der Reaktionsprodukte, als emittierte γ-Strahlung oder als Anregungsenergie der Produkte in Erscheinung. Bei einem negativen Q-Wert ist zum Auslösen der Reaktion eine Mindestenergie E_s in Form von kinetischer Energie des Geschoßteilchens a zuzuführen. Wegen der Impulserhaltung kann hiervon jedoch nur ein Teil für die Kernreaktion wirksam werden, der Rest wird für die Bewegung des gemeinsamen Schwerpunkts der Reaktionsprodukte benötigt: $E_s = -Q(1 + m/M)$ (m Masse des Geschoßteilchens, M Masse des Targetkerns).

3.4.2.2. Wirkungsquerschnitte. In einer Kernreaktion sind nur ein- und auslaufende Teilchen und evtl. der durch einen Rückstoßimpuls beschleunigte Restkern sowie die Strahlung der Folgeprodukte direkt beobachtbar. Bei gleichen Anfangsparametern der stoßenden Partner können unter Umständen verschiedene Reaktionen (↑ 3.4.2.1.) eintreten, d. h. die Endprodukte unterscheiden sich hinsichtlich ihrer Art, der Energien, der Austrittswinkel und ihrer Polarisation (falls die Teilchen einen Spin besitzen). Die quantitative Angabe der Ausbeute erfolgt durch den Wirkungsquerschnitt (s. a. 2.4.3.5.). Er ist definiert als die effektive Fläche, welche der Targetkern dem Strom der Geschoßteilchen für die betrachtete Reaktion bzw. den speziellen Reaktionsablauf bietet:

$$\sigma = \text{Strahlquerschnitt} \cdot \frac{\text{Zahl der Reaktionen pro Kern}}{\text{Zahl der im gesamten Strahlquerschnitt auftreffenden Teilchen}}$$

Somit besitzt eine Reaktion von Targetkern mit Geschoßteilchen mehrere Wirkungsquerschnitte, je nachdem, ob z. B. Streuung, Teilcheneinfang oder Teilchentransfer betrachtet wird. Diese verschiedenen Teilquerschnitte addieren sich zum Gesamtwirkungsquerschnitt für die betrachtete Kombination der Stoßpartner

$$\sigma_{tot} = \sigma_{el} + \sigma_{inel} + \sigma_{einfang} + \sigma_{transf} + \dots$$

Da der Reaktionsquerschnitt häufig mit dem geometrischen Querschnitt des Kerns vergleichbar ist, wird als Einheit 1 barn = 10^{-28} m² benutzt. Tatsächlich können jedoch die Reaktionsquerschnitte um mehrere Größenordnungen über oder unter diesem geometrischen Querschnitt liegen.

Ein Teilchenstrahl erfährt beim Durchgang einer Schichtdicke x des Targets eine exponentielle Abnahme seiner Intensität (↑ 2.4.3.5.). Aufgrund des meist geringen Wirkungsquerschnitts für Kernreaktionen verlieren geladene Geschoßteilchen in einem Target ihre kinetische Energie durch sukzessive Abbremsung an den Hüllenelektronen. Somit kann eine entsprechende Kernreaktion, für die eine Schwellenergie erforderlich ist, nur in einer hinreichend dünnen Materialschicht untersucht werden.

Die Abhängigkeit des Wirkungsquerschnitts von der Energie dei Stoßpartner wird als Anregungsfunktion bezeichnet. Sie kann durch eine Vielzahl schmaler Resonanzen, aber auch durch einen weitgehend glatten Verlauf gekennzeichnet sein.

3.4.2.3. Modelle für Kernreaktionen. Das Verständnis des Mechanismus einer Kernreaktion ist eng mit der Kenntnis der Struktur der Atomkerne verbunden. Die Wechselwirkungen zwischen den Nukleonen erzeugen innerhalb des Kerns ein mittleres Potential (Abb. 3.4.2-1), in

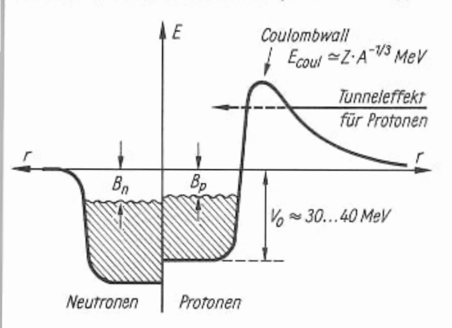


Abb. 3.4.2-1 Mittleres Potential eines Kerns mit der Massenzahl A und der Ladungszahl Z für ein einzelnes Neutron bzw. Proton (*gestrichelt* bereits besetzte Energieniveaus); B_n, B_p Bindungsenergie für ein Neutron bzw. Proton; ein Proton „tunnelt" durch den Coulombwall

dem sich die Nukleonen weitgehend unabhängig voneinander bewegen (↑ 3.3.3.1.). Ein Geschoßteilchen, welches den Atomkern trifft, tritt streng genommen mit der Gesamtheit der Nukleonen in Wechselwirkung. Wegen der Unmöglichkeit der exakten Beschreibung dieses Vorgangs werden vereinfachende Modellvorstellungen benutzt, deren Gültigkeit durch Vergleich mit dem Experiment nachzuweisen ist.

3.4.2.3.1. *Direkte Reaktionen.* Die einfachste Modellbeschreibung ist die der Streuung des Geschoßteilchens am mittleren Potential des Kerns. Sie gilt für den Fall, daß eine Absorption nicht oder kaum eintritt, kann jedoch auch bei komplizierten Reaktionen formal als einer der Beiträge zum Wirkungsquerschnitt betrachtet werden (Potentialstreuung). Die Wechselwirkung des Geschoßteilchens mit nur einem oder wenigen Nukleonen des Targetkerns ist eine weitere einfache Modellvorstellung. Sie ist anwendbar für viele unelastische Streuprozesse und Transferreaktionen. Alle diese Prozesse werden als direkte Reaktionen bezeichnet, da sie ohne Zwischenschritte erfolgen. Ihre Zeitdauer entspricht etwa der Zeit, welche das Geschoßteilchen zum Passieren des Targetkerns benötigt ($\approx 10^{-22}$ s).

3.4.2.3.2. *Compoundkernmodell.* Bei einer anderen Klasse von Reaktionen verteilt sich die verfügbare Energie rasch unter die Nukleonen des Kerns. Es bildet sich ein angeregter Zwischenzustand (Compoundzustand). Dieses Modell wurde von Bohr 1935 zur Interpretation der von Fermi und Amaldi gefundenen schmalen Resonanzen in den Wirkungsquerschnitten langsamer Neutronen vorgeschlagen. Die Resonanzen im Querschnitt entsprechen quasistationären Zuständen des Produktkerns, wobei die Breite ΔE der Resonanzen und die Lebensdauer τ der Zustände durch die Relation $\tau \Delta E \approx \hbar$ verknüpft sind (↑ 3.3.2.3.2.). Einer Resonanzbreite von 1 eV entspricht die Lebensdauer des Zwischenzustands von $6 \cdot 10^{-16}$ s. Die relativ große Lebensdauer des Zwischenzustands garantiert eine weitgehende Unabhängigkeit seines weiteren Schicksals (Zerfall, Spaltung usw.) von der Art und Weise der Bildung. Direkte Reaktionen und Compoundreaktionen sind extreme Beschreibungen einer Kernreaktion. So wird die elastische Streuung nicht ausschließlich durch die Wechselwirkung mit dem mittleren Potential des Targetkerns hervorgerufen, sondern verläuft auch z. T. über einen Compoundzustand.

3.4.2.3.3. *Statistisches Modell.* Die Bohrsche Hypothese erwies sich als äußerst fruchtbar. Sie legte nahe, den Compoundkern mit den Methoden der statistischen Thermodynamik zu beschreiben. Die Ergebnisse des statistischen Modells stimmen in solchen Fällen befriedigend mit den experimentellen Werten überein, in denen es sich um Grobeffekte des Kerns, wie etwa das Verhalten der Niveaudichte, handelt. Der Zerfall des Compoundsystems kann in grober Näherung ebenfalls statistisch als Verdampfung von Nukleonen aus einer Kernmaterie hoher Temperatur behandelt werden. Das weitere Schicksal des Compoundzustands hängt dabei von seinen Zerfalls- bzw. Abregungsmöglichkeiten ab. Im allgemeinen wird sich

der Compoundzustand durch Emission von γ-Strahlung auf den Grundzustand des Zwischenkerns abregen. Jedoch tritt auch die Erscheinung der Kernspaltung und der Teilchenemission aus dem angeregten Zustand des Zwischenkerns auf. Ist der Grundzustand instabil, erfolgt daraufhin radioaktiver Zerfall.

3.4.2.3.4. *Optisches Modell.* Das Modell benutzt ein phänomenologisches Potential anstelle des Targetkerns, dessen Realteil ähnlich dem mittleren Potential aller Nukleonen des Targetkerns die elastische Streuung beschreibt, während ein zusätzlicher Imaginärteil die Absorption des wechselwirkenden Teilchens berücksichtigt. Dieser sehr einfache Potentialansatz läßt die individuelle Struktur der Kerne und somit ihr Resonanzverhalten unberücksichtigt. Es beschreibt dessenungeachtet sehr gut das über die Einzelresonanzen gemittelte Verhalten der Anregungsfunktionen von mittelschweren und schweren Kernen sowie ihren glatten Verlauf bei höheren Anregungsenergien (s. a. 3.3.3.3.).

3.4.2.4. Kernreaktionen und Kernstruktur. Bei der Aufklärung der Struktur von Atomkernen haben die Untersuchungen der Kernreaktionen einen wesentlichen Anteil. Resonanzen von Compoundreaktionen entsprechen Energieniveaus der Produktkerne (Abb. 3.4.2-2). Übergangswahrscheinlichkeiten, Lebensdauern und Winkelverteilungen gestatten Rückschlüsse auf die Parameter der Niveaus (Spin, Isospin, Parität). Darüber hinaus hat sich die Verteilung und Struktur der Resonanzen als wichtiges Hilfsmittel bei der Erforschung der Kernstruktur erwiesen.

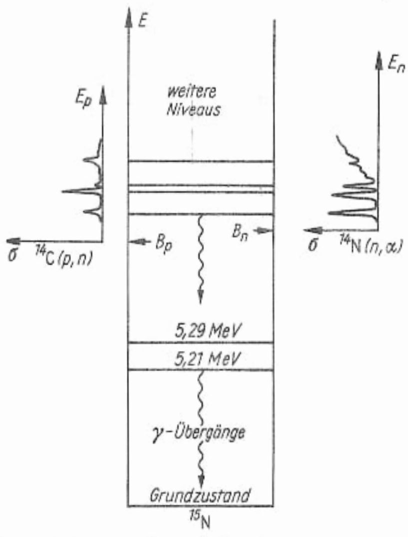


Abb. 3.4.2-2 Schematischer Zusammenhang von Kernstruktur (angeregte Niveaus) und Compoundresonanzen in den Teilchenkanälen (stark vereinfacht; Erläuterung im Text)

Transferreaktionen liefern dagegen direkte Informationen über Niveaus des Targetkerns bzw. des Produktkerns der Reaktion. Das ist durch die Tatsache zu erklären, daß hier Nukleonen direkt aus ihren diskreten Quantenzuständen im Targetkern entfernt werden bzw. in solche diskreten Zustände des Endkerns eingebaut werden. In der unelastischen Streuung werden ebenfalls diskrete Niveaus des Targetkerns angeregt.

3.4.2.5. Spezielle Kernreaktionen. Kernreaktionen sind bezüglich ihrer Mechanismen sowie der beteiligten Partner und der Produkte außerordentlich verschiedenartig. Sie führten zur Darstellung und Untersuchung von mehr als 1500 neuen Nukliden.

3.4.2.5.1. *Kernreaktionen mit Neutronen.* Das Verhalten von Kernreaktionen mit Neutronen wird in einem Energiebereich von etwa 10 Größenordnungen studiert. Für niedrige Neutronenenergien (bis zu 100 eV) dominieren (n, γ)-Prozesse. Einem Abfall des totalen Wirkungsquerschnitts proportional zu $1/v$ folgt ein Bereich schmaler isolierter Resonanzen mit z. T. sehr großen Wirkungsquerschnitten. Die großen Querschnitte bei kleinen Neutronengeschwindigkeiten v sind durch den Wellencharakter des Neutrons bedingt. Die entsprechende de-Broglie-Wellenlänge für langsame Neutronen übersteigt um ein Vielfaches die Dimensionen des Targetkerns. v^{-1} ist proportional zur Aufenthaltsdauer im Wechselwirkungsbereich zwischen Neutron und Targetkern. Mit wachsender Neutronenenergie steigt im Resonanzgebiet der Anteil der elastischen und unelastischen Streuung. Die Zahl der Resonanzen und ihre Breite nehmen zu, so daß sie nicht mehr vollständig isoliert erscheinen. Bei noch höheren Energien zeigt sich eine weitgehende Glättung des Wirkungsquerschnitts, wobei nur flache und sehr breite Resonanzerscheinungen übrigbleiben.

Beim Einfang eines Neutrons erhält der Produktkern außer der kinetischen Energie noch die Bindungsenergie des Neutrons von etwa 8 MeV. Im allgemeinen erfolgt die Abregung zum Grundzustand des Produktkerns durch γ-Emission. Der so entstandene isotope Kern (gleiche Ladung, höhere Masse) ist häufig gegenüber β^--Zerfall instabil (↑ 3.3.1.2.). Neutroneneinfang mit anschließender Teilchenemission ist meist endotherm und setzt somit nur oberhalb einer Schwellenenergie ein. Demgegenüber sind die Reaktionen Li-6 (n, α) H-3 und B-10 (n, α) Li-7 ($Q = 4{,}6$ bzw. 2,78 MeV) exotherm und werden zum Nachweis thermischer Neutronen verwendet.

In schweren Kernen führt der Neutroneneinfang zusätzlich zur Kernspaltung (↑ 3.4.1.). Diese wurde 1938 durch Hahn und Strass-

MANN nachgewiesen. Dabei reicht bei den Targetkernen U-233, U-235, Pu-239, Pu-241 und bei weiteren noch schwereren Kernen bereits die Neutronenbindungsenergie aus, um den Spaltprozeß herbeizuführen. Alle anderen schweren Kerne benötigen dazu noch eine zusätzliche Anregung durch eine entsprechende kinetische Energie des eingefangenen Neutrons.

3.4.2.5.2. *Kernreaktionen mit Protonen.* Der Mechanismus und das Verhalten ist mit dem der Neutronenreaktionen vergleichbar. Es werden Compoundkernreaktionen und direkte Reaktionen sowie ihre Zwischenstufen beobachtet. Die Produkte der (p, γ)-Reaktion (Protoneneinfang) sind im Unterschied zum Neutroneneinfang β^+-Strahler. Der wichtigste Unterschied besteht in der Existenz der Coulombbarriere, welche bei einem Z-fach geladenen Targetkern $E_{coul} \approx (Z/A^{1/3})$ MeV beträgt. Unterhalb dieser Energie ist das Eindringen der Protonen in den Kern nur durch den quantenmechanischen Tunneleffekt (↑ Abb. 3.4.2-1) möglich, so daß hier die Wirkungsquerschnitte gering sind. Die in den Sternen ablaufenden Fusionsreaktionen erfolgen ausschließlich bei kinetischen Energien unterhalb der Coulombbarriere der Stoßpartner. Die geringen Wirkungsquerschnitte bedingen einen entsprechend langsamen Verlauf der Reaktionen und erklären somit die lange Lebensdauer der Sterne.

3.4.2.5.3. *Kernreaktionen mit Deuteronen.* Die Besonderheiten dieser Reaktionen sind durch die geringe Bindungsenergie des Deuterons (2,23 MeV) bedingt. Ein Deuteron kann bereits beim Vorbeifliegen am Targetkern aufbrechen und mit einem seiner Nukleonen eine Reaktion herbeiführen, während das andere Nukleon weiterfliegt, man spricht von der Strippingreaktion (d,p), (d,n). Als Beispiele seien die Reaktionen

d + t → He-4 + n + 17,6 MeV
d + d → He-3 + n + 3,38 MeV
d + d → He-3 + p + 4,0 MeV

aufgeführt, welche die Grundlage für eine mögliche Kernenergiegewinnung auf der Basis der Fusion leichter Kerne (↑ 9.3.6.) darstellen.

3.4.2.5.4. *Kernreaktionen mit α-Teilchen.* Die erste künstliche Kernreaktion wurde 1919 durch RUTHERFORD herbeigeführt: N-14 (α,p) O-17. Dabei wurden die α-Teilchen aus dem radioaktiven Zerfall des RaC′ benutzt. Da bei α-Teilchen gegenüber Protonen eine höhere Coulombschwelle überwunden werden muß, benötigt man für diese Reaktion höhere Geschoßenergien. Stellvertretend genannt sei die Reaktion Be-9 (α,n) C-12, welche lange Zeit in Verbindung mit natürlichen oder künstlichen α-Strahlern (Ra, Rn, Po, Pu, Am) die wichtigste Neutronenquelle darstellte. An dieser Reaktion entdeckte CHADWICK 1932 das Neutron.

3.4.2.5.5. *Kernreaktionen mit schweren Ionen.* Unter schweren Ionen versteht man alle Ionen schwerer als He-4. Ihre Wechselwirkung untereinander ist charakterisiert durch die entsprechend starke Coulombwechselwirkung; die kleine de-Broglie-Wellenlänge relativ zu den Dimensionen des Systems, welche es gestattet, die Bahnen der wechselwirkenden Teilchen bei den hier interessierenden Energien als klassisch zu betrachten; die hohe Drehimpulsübertragung (bis zu 100 $\hbar$) und den Transfer einer größeren Zahl von Nukleonen.

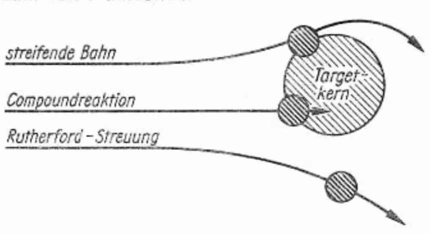


Abb. 3.4.2-3 Verschiedene Möglichkeiten für Schwerionenreaktionen

Entsprechend dem Abstand, unter dem die Reaktionspartner aufeinanderstoßen, unterscheidet man grob drei Bereiche (Abb. 3.4.2-3): 1. die Compoundkernbildung, 2. die streifende Wechselwirkung bzw. den Transfer von Nukleonengruppen und 3. die Rutherfordstreuung, bei welcher Coulombanregung und Tunnelung einzelner Nukleonen eintreten kann (Nukleonentransfer). *Compoundkernreaktionen* treten gegenüber den Reaktionen mit leichten Geschossen noch bei wesentlich höheren Energien auf, da sich die kinetische Energie einer Vielzahl von relativ niederenergetischen Nukleonen sehr schnell auf alle Nukleonen des Systems verteilen kann. Die mit schweren Ionen erzeugbaren Reaktionsprodukte sind außerordentlich vielfältig, lassen sich jedoch in den allgemeinen Rahmen des bisher Behandelten einordnen.

Reaktionen mit schweren Ionen werden für die Erzeugung und Untersuchung von Transurannukliden verwendet. Der Vorstoß über das auf der Erde vorkommende Gebiet der Elemente hinaus ist für das Verständnis der Kernstruktur, aber auch für die Lösung kosmischer und kosmologischer Fragestellungen sehr wichtig. Je höher die Ladungszahl eines Nuklids oberhalb von Uran ist, desto kürzer ist die Lebensdauer seiner Isotope gegenüber spontaner Spaltung, α- und β-Zerfall, und desto geringer ist die Wahrscheinlichkeit ihrer Bildung in einer Compoundkernreaktion. Bisher gelang die Darstellung von Nukliden bis zur Ordnungszahl 107. Von besonderem Interesse ist die Vorhersage der Kerntheoretiker, daß in der Nähe der Ladungszahl 114 eine „stabile Insel" existieren

soll, hier also Nuklide mit relativ großen Lebensdauern existieren könnten. Die Suche nach Nukliden der stabilen Insel erfolgt derzeitig sowohl in Kernreaktionen mit immer schwereren Reaktionspartnern als auch in der Natur, wo in bestimmten Mineralien bzw. in Meteoriten Restmengen der Isotope der stabilen Insel existieren sollten, falls ihre Lebensdauer mit dem Alter des Sonnensystems vergleichbar wäre. Eine der Möglichkeiten zur Erzeugung des Elements mit der Ordnungszahl 114 stellt die Reaktion $^{244}_{94}Pu + ^{48}_{20}Ca \rightarrow ^{288}_{114}X + 4n$ dar, wozu die Calciumatomkerne auf Energien von etwa 300 MeV beschleunigt werden müssen.

3.4.2.5.6. *Fotokernreaktionen.* Durch die Absorption von γ-Quanten können Kerne so weit angeregt werden, daß die Energie für die Emission von Teilchen oder für deren Spaltung ausreicht. Ein solcher Kernfotoeffekt wurde bereits 1934 von CHADWICK und GOLDHABER bei der Zerlegung des Deuteriums beobachtet und ermöglichte die erste genaue Massenbestimmung des Neutrons. (γ,n)-Prozesse werden für die Erzeugung hoher Neutronenflüsse an Teilchenbeschleunigern benutzt (↑ 3.4.4.).

3.4.2.5.7. *Kernreaktionen mit Teilchen höchster Energien.* Entsprechend den vielen Freiheitsgraden, welche ein hochenergetisches Geschoßteilchen im Targetkern anregen kann, sind verschiedenartige Reaktionen möglich: Die große übertragene Energie kann sich in einer kollektiven Anregung äußern, welche zur Verdampfung vieler Nukleonen und zur Spaltung des Systems in mehrere Bruchstücke führt; man spricht von *Spallation.* Es kann aber auch der Fall eintreten, daß ein hochenergetisches Nukleon auf seinem Wege durch den Kern nur mit einzelnen Nukleonen reagiert. Dabei kommt es zum Herausschlagen dieser Nukleonen, ohne den Restkern wesentlich zu beeinflussen, was als *knockout-Prozeß* bezeichnet wird. Weitere Effekte sind die *Bildung neuer Teilchen* wie Mesonen, Nukleonen und Hyperonen, sofern die entsprechenden Energieschwellen überschritten sind.

Eine Besonderheit wird beim Stoß von schweren Ionen mit relativistischen Energien erwartet: Wird die Relativbewegung der Reaktionspartner größer als die Schallgeschwindigkeit in der Kernmaterie, sollte sich eine Stoßwelle mit erhöhter Nukleonendichte bilden. Die hierbei beobachteten Effekte könnten Aufschluß über Fragen der *überdichten Materie* und evtl. stabiler Bereiche überdichter Materie geben.

3.4.3. Angewandte Radioaktivität

3.4.3.1. Gesetzmäßigkeiten des radioaktiven Zerfalls. Von den rund 1 500 bis heute bekannten Kernarten sind nur knapp 270 stabil, während sich die übrigen, eventuell über andere instabile Zwischenstufen, in stabile Kerne umwandeln. Wegen der dabei ausgesendeten verschiedenartigen Strahlen bezeichnet man instabile Kerne als *radioaktiv* und die gesamte Erscheinung der spontanen Kernumwandlung als *radioaktiven Zerfall.*

Im allgemeinen sind die radioaktiven Kerne mit inaktiven Kernen des gleichen oder anderer Elemente vermischt. Nur in trägerfreien Präparaten ist das Radionuklid durch chemische Verfahren oder durch Isotopentrennung von dem inaktiven *Trägermaterial* abgetrennt. Die Mehrzahl der Radionuklide kann nur künstlich aus Kernreaktionen erhalten werden. In der Natur findet man nur die langlebigen radioaktiven Kerne der schwersten Elemente Thorium und Uranium mit ihren kurzlebigen Tochterprodukten, einige andere Radionuklide sehr großer Lebensdauer, wie K-40, sowie sehr kleine Mengen einiger Radionuklide, die durch Kernreaktionen in der Atmosphäre unter dem Einfluß der *kosmischen Strahlen* gebildet werden; z. B. C-14 und Tritium.

3.4.3.1.1. *Statistischer Charakter und Zerfallsgesetz.* Der radioaktive Zerfall ist *statistischer Natur*, die Kerne wandeln sich unabhängig voneinander mit einer vorgegebenen Zerfallswahrscheinlichkeit um. Die *Zerfallswahrscheinlichkeit* bestimmt die *mittlere Lebensdauer* der Kerne des Nuklids. Der Zeitpunkt, in dem sich ein bestimmter Kern umwandelt, ist damit nicht festgelegt. Die Zerfallswahrscheinlichkeit der Kerne ist konstant und unabhängig von der Vorgeschichte. Zwei Kerne des gleichen Radionuklids, von denen der eine soeben entstanden ist, während der andere aus einer früheren Bestrahlung stammt, wandeln sich mit der gleichen Wahrscheinlichkeit innerhalb der nächsten 24 Stunden um.

Aus der konstanten Zerfallswahrscheinlichkeit und der Unabhängigkeit der einzelnen Zerfallsakte voneinander folgt, daß der Bruchteil dn der Atomkerne eines Nuklids, die sich in einem kleinen Zeitintervall dt umwandeln, proportional zur Anzahl n der jeweils vorhandenen radioaktiven Kerne ist. Der Proportionalitätsfaktor λ wird als Zerfallskonstante bezeichnet, und es gilt $dn/dt = -\lambda n$. Durch Integration dieser Differentialgleichung erhält man $n = n_0 \exp(-\lambda t)$, wenn n_0 die Anzahl der Kerne zur Zeit $t = 0$ bedeutet.

Man benutzt als bequemes und anschauliches Maß für die Lebensdauer eines Radionuklids die *Halbwertszeit* $t_{1/2} = \ln 2/\lambda$; das ist die Zeitspanne, in der sich die Hälfte der vorhandenen Atomkerne umwandelt. Tritium hat eine Halbwertszeit von 12,26 Jahren; von 10^6 Tritiumatomen werden daher nach rund 12,5 Jahren nur noch rund die Hälfte, nach 25 Jahren ein Viertel, nach 37,5 Jahren ein Achtel und nach

50 Jahren ein Sechzehntel der ursprünglichen Anzahl vorhanden sein. Durch Auftragen der Kernzahl über der Zeit ergibt sich der charakteristische Exponentialverlauf (Abb. 3.4.3-1).

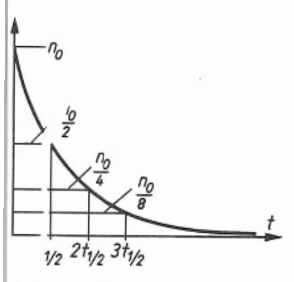


Abb. 3.4.3-1 Exponentieller Verlauf des radioaktiven Zerfalls

Die Halbwertszeiten aller bekannten Radionuklide liegen im Bereich zwischen 10^{-10} s und 10^{1} a. Am häufigsten sind allerdings Radionuklide mit einer Halbwertszeit von einigen Minuten bis zu einigen Jahren.

3.4.3.1.2. *Einfache Anwendungen des Zerfallsgesetzes.* Aus dem Zerfallsgesetz lassen sich die Zeitgesetze für die Bildung von Radionukliden sowie für den Zerfall von radioaktiven Folgeprodukten ableiten.

3.4.3.1.2.1. *Zeitgesetz für die Bildung eines Radionuklids.* Künstliche Radionuklide werden durch Bestrahlung von Targets im Teilchenstrom eines Reaktors oder eines Beschleunigers hergestellt. Während der Bestrahlungszeit kann die Teilchenstromdichte Φ als konstant vorausgesetzt und die kleine relative Änderung der Konzentration der beschossenen Kerne N vernachlässigt werden. Daher ist die zu Φ proportionale *Bildungsrate r*, d. h. die Anzahl der in der Zeiteinheit *entstehenden Kerne*, konstant. Andererseits *zerfällt* bereits während der Bestrahlung ein Teil der gebildeten radioaktiven Kerne. Insgesamt lautet die Beziehung für das *Anwachsen* der Anzahl radioaktiver Kerne im Target somit $dn/dt = r - \lambda n$. Die Lösung dieser Gleichung mit den Anfangsbedingungen $n = 0$ bei $t = 0$ liefert $n = r\lambda[1 - \exp(-\lambda t)]$. Wie der Verlauf dieser Funktion erkennen läßt, nähert sich die Aktivität nach einigen Halbwertszeiten einem Sättigungswert. Danach bleibt die Aktivität konstant, da ebensoviele Kerne zerfallen wie neu gebildet werden. Um kurzlebige Radionuklide herzustellen, genügen daher kleine Bestrahlungszeiten.

Andererseits läßt sich bei der Herstellung langlebiger Kerne in vernünftigen Bestrahlungszeiträumen nur ein kleiner Bruchteil der Sättigungsaktivität erzielen.

3.4.3.1.2.2. *Zerfall mit radioaktivem Folgeprodukt.* In vielen Fällen bildet sich als Produkt einer radioaktiven Umwandlung wieder ein Radionuklid. Aus den natürlichen Strahlern Uranium und Thorium leiten sich ganze *Zerfallsreihen* instabiler Nuklide ab, bevor als stabile Endprodukte Blei- bzw. Bismutkerne entstehen. Auch bei künstlich hergestellten Radionukliden werden *radioaktive Folgeprodukte* beobachtet. Liegt zur Zeit $t = 0$ in der Probe nur die Muttersubstanz 1 vor, so steigt zunächst die Konzentration des Tochterprodukts 2 an, strebt aber im Falle $(t_{1/2})_2 \ll (t_{1/2})_1$ einem Grenzwert zu, bei dem sich die Bildungs- und die Zerfallsrate genau die Waage halten. Zwischen Mutter- und Tochternuklid stellt sich ein radioaktives Gleichgewicht ein. Nach dem Zerfallsgesetz gilt $dn_1/dt = -\lambda_1 n_1$ und $dn_2/dt = \lambda_1 n_1 - \lambda_2 n_2$. Da im Gleichgewichtsfall $dn_2/dt = 0$, folgt $n_1/n_2 = \lambda_2/\lambda_1 = (t_{1/2})_1/(t_{1/2})_2$. Im Fall des radioaktiven Gleichgewichts stehen die Mengen der beiden Kernarten demnach im gleichen Verhältnis wie die Halbwertszeiten.

Ist dagegen die Halbwertszeit des Mutternuklids nicht sehr viel größer als die des Tochternuklids, so kann auch die Konzentration des Mutternuklids nicht mehr als konstant betrachtet werden. Falls immer noch $(t_{1/2})_1 > (t_{1/2})_2$ gilt, klingen Mutter- und Tochternuklid nach Einstellung eines konstanten Konzentrationsverhältnisses gemeinsam mit der Halbwertszeit $(t_{1/2})_1$ ab, während sich für den Fall $(t_{1/2})_1 < (t_{1/2})_2$ gar kein stationärer Zustand ausbilden kann.

3.4.3.2. Wechselwirkung radioaktiver Strahlung mit Materie

3.4.3.2.1. *Absorption und Strahlenwirkungen.* Radioaktive Strahlen werden beim Durchgang durch Stoffe aus ihrer Richtung abgelenkt oder *gestreut* und in ihrer Intensität geschwächt oder *absorbiert*. Sie übertragen dabei die gesamte oder einen Teil ihrer Energie auf die getroffenen Atome. Wegen der sehr kleinen Kernabmessungen beschränkt sich diese Wechselwirkung hauptsächlich auf die Atomhüllen; die sehr seltenen „Kerntreffer" können bei der Betrachtung von Absorptionserscheinungen fast stets vernachlässigt werden. Im Verlauf der Wechselwirkung werden die Strahlenteilchen abgebremst und schließlich absorbiert, d. h., sie hören entweder auf zu bestehen, wie γ-Quanten oder Positronen, oder sie kommen praktisch zur Ruhe und werden in die Absorberstoffe eingebaut. Die *Stärke der Absorption* und der *Streuung* hängen von der Art und von der Energie der Strahlung sowie von der Dichte und von der Zusammensetzung des absorbierenden Materials ab. Während α-Strahlen bereits von dünnen Aluminium- oder Plastfolien vollständig absorbiert werden, können energiereiche β-Strahlen noch durch Plastmaterial von einigen Millimetern Stärke und γ-Strahlen sogar durch mehrere Zentimeter dicke Eisen- oder Blei-

schichten teilweise hindurchtreten. Im allgemeinen gilt, daß energiereiche Strahlen durchdringungsfähiger sind als energieärmere; man bezeichnet deshalb energiereiche Strahlen etwa oberhalb 0,5 MeV als *harte Strahlen* und energiearme Strahlen als *weiche*.

Absorption und Streuung bilden eine wichtige Grundlage für *Strahlenschutzmaßnahmen* sowie für zahlreiche *technische Anwendungen* der Radioaktivität.

Die mannigfachen Veränderungen in den bestrahlten Stoffen, die *Strahlenwirkungen*, sind eine Folge der Energie, die von den Strahlenteilchen auf die Atome übertragen wird, und äußern sich als physikalische, chemische oder biologische Veränderungen der Stoffe. Primär werden meist ein oder mehrere Hüllenelektronen aus den Atomen herausgeschlagen oder zumindest in höhere Energiezustände angeregt. Als Folge davon werden chemische Bindungen gelöst, und es bilden sich geladene und neutrale Molekülbruchstücke, d. h. *Radikale*, die ihrerseits vielseitige sekundäre chemische Reaktionen eingehen können. In belebter Substanz geben diese chemischen Veränderungen wiederum zu biologischen Wirkungen Anlaß. Die Strahlenwirkungen werden vielseitig angewendet (s. a. 3.4.3.7.5.), insbesondere bilden sie die Grundlage aller *Nachweisverfahren* für radioaktive Strahlen.

3.4.3.2.2. *Durchgang von α-Teilchen durch Materie.* α-Teilchen geben ihre Energie beim Durchgang durch Materie fast ausschließlich durch die elektrostatische Wechselwirkung an die Elektronen der Atomhüllen ab. Die sehr seltenen Kernstöße oder gar Kernreaktionen tragen zum gesamten Energieverlust praktisch nichts bei. Die Wechselwirkung kann als „Stoß" des α-Teilchens mit Hüllenelektronen aufgefaßt werden. Aus dem Impuls- und dem Energiesatz folgt dann, daß es in einem einzelnen Stoßprozeß wegen seiner viel größeren Masse nur einen kleinen Bruchteil seiner Bewegungsenergie an ein Elektron übertragen kann. Die Impulsänderung ist sehr gering, das α-Teilchen wird kaum aus seiner Bahn abgelenkt und hinterläßt in der Nebelkammer oder Fotoschicht geradlinige Bahnen.

Als Folge der Wechselwirkung werden Elektronen aus der Atomhülle herausgeschlagen. Im Durchschnitt beträgt die von einem α-Teilchen an ein Elektron übertragene Energie 100 bis 200 eV; einige Elektronen erhalten sogar eine Energie von 1 keV. Diese Elektronen fliegen dann selbst ein Stück durch die Materie und ionisieren sekundär einige weitere Atome in der Umgebung der Bahn des α-Teilchens. Es bilden sich *Ionenpaare*, die bei Edelgasen aus einem Atomion und dem abgetrennten Elektron bestehen. Bei Absorbern, die aus Molekülen aufgebaut sind, werden daneben auch negative Ionen, Dissoziationsprodukte u. a. beobachtet. Die Anzahl der gebildeten Ionenpaare hängt von der Energie des α-Teilchens und vom durchstrahlten Material ab. In der Luft erzeugt ein α-Teilchen bei vollständiger Abbremsung ein Ionenpaar je 35 eV seiner Energie. Danach hinterläßt ein α-Teilchen des Radiums mit $E_\alpha = 4{,}777$ MeV längs seines Weges in Luft rund 140000 Ionenpaare. Als *spezifische Ionisation* eines Strahlenteilchens definiert man die Anzahl der Ionenpaare, die von dem Teilchen in Luft unter Normalbedingungen pro Millimeter Weglänge gebildet werden.

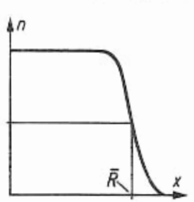


Abb. 3.4.3-2 Reichweite von α-Teilchen in Luft; $\bar{R}$ mittlere Reichweite, *n* Anzahl der Teilchen

Monoenergetische α-Strahlen haben infolge der zahlreichen zum Abbremsen notwendigen Stoßprozesse eine annähernd einheitliche Reichweite. In einem Strahlenbündel einheitlicher Anfangsenergie ergibt sich für die Anzahl der nach einem bestimmten Luftweg noch nachzuweisenden Teilchen (Abb. 3.4.3-2) die *Braggsche Kurve*. Sind mehrere Gruppen von α-Strahlen mit unterschiedlicher Energie vorhanden, so weist die Absorptionskurve mehrere ziemlich scharfe Sprünge auf. Die abgebremsten α-Teilchen fangen Elektronen ein und verwandeln sich damit in Heliumatome, die in kompakten Materialien eingeschlossen bleiben.

3.4.3.2.3. *Durchgang von β-Strahlen durch Materie.* Ebenso wie bei α-Teilchen erfolgt das Abbremsen von β-Strahlen fast ausschließlich durch elektrostatische Wechselwirkung oder „Stöße" mit den Hüllenelektronen. Wesentliche Unterschiede gegenüber den α-Strahlen ergeben sich jedoch hinsichtlich der Streuwinkel und der Ionisierungsdichte. Da die β-Teilchen eine viel kleinere Masse als die α-Teilchen haben, werden sie bei der Wechselwirkung mit den Hüllenelektronen selbst viel stärker aus ihrer Bahn abgelenkt, so daß ihre Bahnen nicht geradlinig verlaufen, sondern starke Krümmungen aufweisen. Die Reichweite der β-Strahlen ist bei gleicher Energie viel größer als die von α-Strahlen, und die spezifische Ionisation, die bei 8-MeV-Elektronen nur 6,5 Ionenpaare/mm beträgt, ist entsprechend geringer.

Elektronen einheitlicher Energie sollten ebenso wie α-Teilchen eine einheitliche Reichweite ha-

ben. Wegen der starken Richtungsänderungen zeigen die längs einer Geraden gemessenen effektiven Reichweiten jedoch starke Schwankungen.
Positronen und Elektronen verhalten sich bei der Wechselwirkung mit Materie nahezu gleichartig. Abgebremste Positronen zerstrahlen jedoch zusammen mit einem Hüllenelektron der durchstrahlten Materie. Durch Vernichtung der beiden Teilchen wird ihre Ruhenergie in Form von zwei in entgegengesetzter Richtung auseinanderfliegenden Quanten von je 0,51 MeV abgestrahlt. Infolgedessen wird bei Positronenstrahlen stets eine Quantenstrahlung dieser Energie beobachtet, die sogenannte *Vernichtungsstrahlung*.
Bei der Absorption von β^+- und β^--Strahlen kommt es schließlich zur Entstehung eines kontinuierlichen Untergrunds von *Bremsstrahlung*. Die Intensität dieser Quantenstrahlung nimmt proportional zur Ordnungszahl des Absorbers zu.

3.4.3.2.4. *Durchgang von γ-Strahlen durch Materie*. Die Wechselwirkung von γ-Strahlen mit Materie verläuft wesentlich anders als die von α- und β-Strahlen. Während geladene Teilchen ihre Energie portionsweise in vielen Einzelprozessen an die Hüllenelektronen übertragen, geben die γ-Quanten in einem Akt entweder ihre gesamte Energie oder zumindest einen großen Bruchteil an ein Hüllenelektron ab. Die Wahrscheinlichkeit dafür, daß überhaupt eine Wechselwirkung zustande kommt, ist allerdings klein, und deshalb sind γ-Strahlen viel durchdringungsfähiger als Strahlen geladener Teilchen.
In der Nebelkammer oder in Fotoschichten hinterlassen die γ-Quanten selbst keine Bahnspuren, doch beobachtet man die Spuren energiereicher Elektronen, die bei der Wechselwirkung von Quanten mit Materie gebildet wurden. Die Intensität der γ-Strahlen nimmt entsprechend $\exp(-\mu x)$ exponentiell mit wachsender Dicke x der durchstrahlten Schicht ab. μ ist der Absorptionskoeffizient.
Die Absorption wird durch drei voneinander unabhängige Effekte hervorgerufen. Dementsprechend setzt sich der Absorptionskoeffizient $\mu = \mu_{Co} + \mu_{Ph} + \mu_{Pa}$ als Summe der entsprechenden Schwächungskoeffizienten zusammen; dabei beschreiben μ_{Co} die Absorption durch den *Comptoneffekt*, μ_{Ph} die Absorption durch den *Fotoeffekt* und μ_{Pa} die Absorption durch den *Paarbildungseffekt*. Beim *Comptoneffekt* gibt das γ-Quant beim inelastischen Stoß mit einem Hüllenelektron einen Teil seiner Energie ab. Der Bruchteil der an das Elektron übertragenen Energie kann je nach dem Stoßwinkel sehr unterschiedlich sein. Die *Comptonelektronen* haben daher ein kontinuierliches Energiespektrum, dessen Maximum etwa bei der halben Quantenenergie liegt. – Beim *Fotoeffekt* wird die gesamte Quantenenergie auf ein Hüllenelektron übertragen. Es entsteht ein energiereiches Elektron, dessen kinetische Energie gleich der Quantenenergie ist, vermindert um die Ablösearbeit des Elektrons. – Bei der *Paarbildung*, die nur bei Quantenenergien oberhalb 1,02 MeV stattfinden kann, entsteht ein Elektron-Positron-Paar. Aus Gründen der Impulserhaltung kann sich die Paarbildung nur in der Nähe eines Atomkerns vollziehen.
Der Beitrag der drei Effekte zur gesamten Absorption von γ-Strahlen hängt in erster Linie von der Energie der Quanten und von der Ordnungszahl des Absorbers ab. Aus dem Verlauf des gesamten Absorptionskoeffizienten sowie der drei Anteile als Funktion der Quantenenergie für Blei als Absorbermaterial (Abb. 3.4.3-3) erkennt man, daß bei niedriger Quantenenergie bis 0,5 MeV der Beitrag des *Fotoeffekts* überwiegt. Am wahrscheinlichsten ist dabei die Ablösung von inneren Elektronen der K- oder L-Schale. Die Absorption durch den Fotoeffekt nimmt mit wachsender Quantenenergie ab. Auch der Beitrag des *Comptoneffekts* sinkt mit zunehmender Energie, und zwar nimmt die Wahrscheinlichkeit für den Comptoneffekt oberhalb 0,5 MeV ungefähr proportional E^{-1} ab. Bei mittlerer Quantenenergie (0,5 bis 4,5 MeV in Blei) liefert der Comptoneffekt den hauptsächlichen Anteil. Ab einer Quantenenergie von etwa 4 MeV steigt die Wahrscheinlichkeit für die Paarbildung proportional zum Logarithmus der Quantenenergie an, so daß die Paarbildung bei sehr hohen Energien schließlich für die Absorption bestimmend wird. Das ist bei den γ-Strahlen der Radionuklide jedoch nicht der Fall, denn die Quantenenergie erreicht selbst bei sehr harten Strahlern nur wenige MeV.
Die Absorption von γ-Strahlen weist für jeden Absorber ein Minimum auf, d. h. einen Energie-

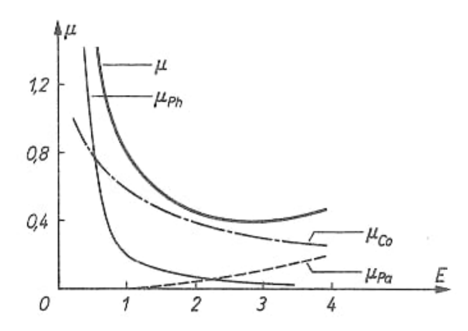


Abb. 3.4.3-3 Absorption von γ-Strahlen in Blei als Funktion der Quantenenergie E in MeV; μ gesamter Absorptionskoeffizient in cm^{-1} Anteile μ_{Ph} des Foto-, μ_{Pa} des Paarbildungs- und μ_{Co} des Comptoneffekts

wert, für den der Stoff eine maximale Durchlässigkeit für γ-Strahlen hat. Für die Lage des Minimums beim Blei erhält man (↑ Abb. 3.4.3-3) einen Wert von rund 3 MeV.

Die Wahrscheinlichkeit aller drei Absorptionsprozesse hängt schließlich auch von der Ordnungszahl des Absorbermaterials ab. Am stärksten ist der Einfluß beim *Fotoeffekt*, dessen Wahrscheinlichkeit mit Z^4 bis Z^5 ansteigt. Der *Comptonanteil* wird dagegen durch die Gesamtzahl der Hüllenelektronen festgelegt. Seine Wahrscheinlichkeit steigt daher proportional mit Z an. Die Wahrscheinlichkeit für die *Paarbildung* wächst etwa proportional Z^2.

3.4.3.3. Nachweis und Messung von Kernstrahlung

3.4.3.3.1. *Strahlungsdetektoren.* Eine Kernstrahlung wird in einem Detektor meist dadurch angezeigt, daß durch sie elektrische Signale entstehen, die in einem nachgeschalteten Strahlungsmeßgerät registriert werden. Die Wirkungsweise der Detektoren beruht darauf, daß die Kernstrahlung die Materie beim Durchgang *ionisiert*, *anregt* oder in ihr *chemische Reaktionen* auslöst. Geladene Teilchen erzeugen diese Wirkungen direkt, Röntgen- und γ-Strahlung indirekt durch *Foto-*, *Compton-* sowie *Paarbildungseffekt* und Neutronen über Kernreaktionen bzw. durch Stoßprozesse, bei denen geladene Teilchen entstehen bzw. in Bewegung gesetzt werden. Im folgenden werden die wichtigsten Detektoren betrachtet. Die *Ionisationskammer* besteht im einfachsten Fall aus einem abgeschlossenen, mit Luft oder Edelgas gefüllten Plattenkondensator (Abb. 3.4.3-4). Im Kammervolumen erzeugt ein geladenes Teilchen *Ionenpaare*, deren Anzahl je Weglängeneinheit von der *Art* und *Energie* des Teilchens sowie von der *Gasart* und dem Gasdruck abhängt. Die Ladungsträger wandern wegen des elektrischen Felds zu den Elektroden und erzeugen einen Kammerstrom I, der mit der Kammerspannung U ansteigt. Bei der Sättigungsspannung U_s von einigen 100 V werden alle gebildeten Ladungsträger von den Elektroden abgesaugt. Unterhalb dieser Spannung besteht zwischen Strom und Spannung ein etwa linearer Zusammenhang (Abb. 3.4.3-5). Hinsichtlich der Betriebsart gibt es zwei Varianten, nämlich Impuls- und Strombetrieb. Im Impulsbetrieb arbeitende Ionisationskammern werden vorzugsweise für stark ionisierende Teilchen mit geringer Reichweite eingesetzt, z. B. α-Teilchen, Protonen oder Spaltprodukte. Im Strombetrieb arbeitende Ionisationskammern werden dagegen vor allem für β- und γ-Strahlung eingesetzt; in der Industrie bei Strahlenschranken, Dickenmeßgeräten u. a. und im Strahlenschutz zur Dosisleistungsmessung.

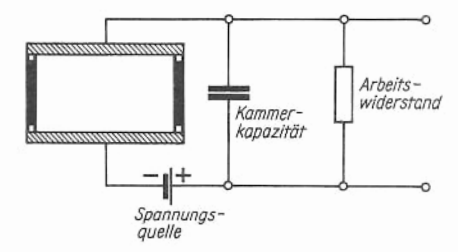


Abb. 3.4.3-4 Schema einer Ionisationskammer

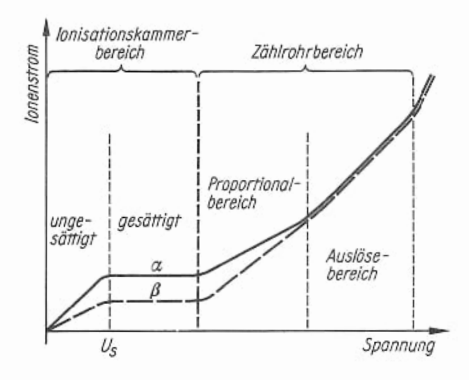


Abb. 3.4.3-5 Ionenstrom als Funktion der Detektorspannung für α- und β-Teilchen

Zum Nachweis von Neutronen wird die Kammer mit Bortrifluoridgas gefüllt oder mit Bor ausgekleidet, wodurch in einer (n,α)-Reaktion α-Strahlen für den Nachweis entstehen. Neben diesen *Borkammern* benutzt man auch *Spaltungskammern*, bei denen die Wände mit reinem U-235 belegt sind.

In einem *Zählrohr* ist die Spannung so hoch, daß die erzeugten Elektronen weitere Atome ionisieren und anregen. Die von den angeregten Atomen emittierten energiereichen Photonen ionisieren ihrerseits wieder und erzeugen eine *Lawine*. Die Entladung greift auf das gesamte Zählrohrvolumen über. Das Zählrohr besteht im einfachsten Fall aus einem Zylinder als Katode, die Anode ist als axialer dünner Draht, häufig aber auch als Schleife ausgebildet. Je nach Höhe der angelegten Spannung wird das Zählrohr in zwei Arbeitsbereichen betrieben, im *Proportionalitäts-* oder im *Auslösebereich*.

Im *Proportionalzählrohr* wird der Strom der primär gebildeten Ionen um einen Faktor 10 bis 10^5 verstärkt. Der entstehende Strom bleibt jedoch zur Primärionisation proportional, so daß eine Energiebestimmung der einfallenden Teilchen möglich ist.

Beim *Auslösezählrohr* oder *Geiger-Müller-Zählrohr* ist die Impulshöhe unabhängig von der Primärionisation und damit von der Energie der einfallenden Teilchen; sie hängt nur von der

angelegten Spannung ab. Das Auslösezählrohr kann daher nur zur Teilchenzählung, nicht aber zur Energiebestimmung der Teilchen benutzt werden. Um eine Dauerentladung im Zählrohr zu verhindern, muß die Entladung durch einen sehr großen Arbeitswiderstand, durch eine spezielle elektronische Schaltung oder durch einen Zusatz von organischen Löschgasen zum Abreißen gebracht werden. In der dazu notwendigen *Totzeit* kann das Zählrohr keine weiteren Teilchen registrieren. Füllgas und Zählrohrgeometrie bestimmen die Totzeit, die für Auslösezählrohre etwa 10^{-4} s, für Proportionalzählrohre rund 10^{-6} s beträgt.

γ-Strahlung wird durch Sekundäreffekte nachgewiesen, die in der Wandung und im Füllgas Elektronen erzeugen; die *Nachweisempfindlichkeit* ist jedoch gering, nur etwa 1%. Für β-Teilchen muß die Zählrohrwandung genügend dünn, oder ein Zählrohrfenster aus dünnem Glimmer vorhanden sein, sogenannte Glocken oder *Fensterzählrohre*.

Bei *Szintillationszählern* wird in bestimmten Leuchtstoffen oder *Szintillatoren* die Energie von Teilchen oder Quanten ganz oder teilweise in Fluoreszenzlicht umgewandelt. Die dabei als Lichtblitze oder Szintillationen ausgesandten Lichtquanten lösen auf der Fotokatode eines *Sekundärelektronenvervielfachers* Fotoelektronen aus, die an weiteren Dynoden *Sekundärelektronen* auslösen, so daß ein millionenfach verstärkter Fotoelektronenstrom als Stromimpuls entsteht (Abb. 3.4.3-6). Hohes Ansprechvermögen für γ-Strahlung, hohes zeitliches Auflösungsvermögen bis zu 10^{-10} s, eine nahezu unbegrenzte Lebensdauer und gute Proportionalität zwischen Teilchen- bzw. Quantenenergie und Stromimpulsamplitude sind die besonderen Vorteile des Szintillationszählers. Als Szintillatoren eignen sich kristalline anorganische und

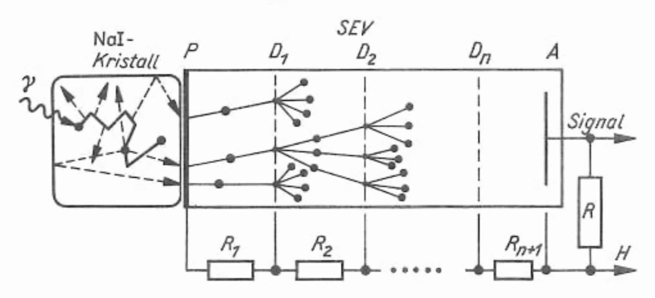


Abb. 3.4.3-6 Schematische Darstellung des Szintillationsdetektors; das γ-Quant wird durch Fotoeffekte im NaI-Kristall und einen Sekundärelektronenvervielfacher (SEV) nachgewiesen; P Fotokatode, D_1 bis D_n Dynoden, A Anode, R_1 bis R_{n+1} Widerstände des Spannungsteilers, R Arbeitswiderstand

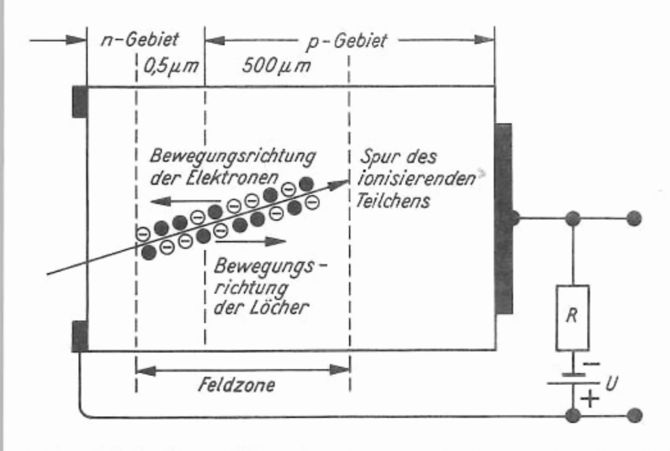


Abb. 3.4.3-7 Schematische Darstellung der Wirkungsweise eines Halbleiterdetektors (*pn*-Zähldiode); ● Defektelektronen (Löcher), ⊖ Elektronen, R Arbeitswiderstand

organische Stoffe, feste und flüssige organische Lösungen sowie Edelgase. Die Auswahl richtet sich nach dem speziellen Verwendungszweck; für den Nachweis schwerer Kerne und Kernteilchen eignen sich Kristallphosphore, z. B. Zinksulfid, am besten, für β-Teilchen organische Substanzen, wie Anthracen und Stilben und für γ-Quanten besonders durch Thalliumzusatz aktivierte Natriumiodidkristalle.

Halbleiterdetektoren sind vergleichbar mit einer Flächendiode großer Ausdehnung. Beim Anlegen einer Spannung in Sperrichtung werden die vorhandenen Raumladungen getrennt, und es kommt zum Aufbau eines inneren elektrischen Felds. In diesem Halbleiterkristall erzeugt ein energiereiches geladenes Teilchen längs seines Wegs durch Ionisation Elektronen-Defektelektronen-Paare, die im Raumladungsgebiet analog zum Vorgang in der Ionisationskammer durch das elektrische Feld getrennt und schnell an die Feldgrenzen befördert werden (Abb. 3.4.3-7). Es entsteht ein kurzzeitiger Stromstoß. Nach der Entladung stellt sich durch Diffusion freier Ladungsträger der alte Zustand schnell wieder her, und der Detektor ist erneut arbeitsbereit. Da die Empfindlichkeit der Halbleiterdetektoren von der Größe der Raumladungszone abhängt, diese aber im allgemeinen klein ist, werden sie vorwiegend zur Messung von stark ionisierenden Teilchen, von Spaltprodukten, Protonen, α- und β-Strahlung benutzt. Auch γ-Strahlung kann nachgewiesen werden, wenn sie durch Compton- oder Fotoeffekt im Halbleiterkristall Elektronen auslöst. Da die zur Erzeugung von Ladungsträgern notwendige Ionisationsenergie wesentlich geringer ist als z. B. bei der Ionisationskammer oder dem Szintillationsdetektor, die Ausbeute an Ladungsträgern demzufolge auch größer, ergibt sich ein sehr gutes energetisches Auflösungsvermögen. Die Nachweisempfindlichkeit ist jedoch aufgrund des geringen Zählvolumens und der niedrigen Ordnungszahl von Silicium gering. Germaniumhalbleiterdetektoren zeigen bessere Ergebnisse, jedoch ist die Eigenleitung bei Zimmertemperatur so hoch, daß derartige Detektoren bei tiefen Temperaturen betrieben werden müssen.

Auch mit Hilfe *fotografischer Emulsionen*, den *Kernspuremulsionen*, gelingt es, den Verlauf der Bahnen einzelner Teilchen aufzuzeichnen. Durch schnelle Elektronen, Mesonen, Ionen und Atome und, infolge von Sekundäreffekten, auch durch Röntgen- und γ-Strahlung sowie Neutronen werden diese geschwärzt. Aus dem Grad der Schwärzung kann auch die Intensität der Strahlung ermittelt werden. *Kernspurplatten* sind besonders geeignet, seltene Ereignisse zu registrieren, z. B. (γ,α)-Reaktionen oder die Strahlung sehr schwacher radioaktiver Substanzen.

3.4.3.3.2. *Strahlungsmeßgeräte.* Dem Detektor (↑ 3.4.3.3.1.) wird die notwendige Spannung aus einer Hochspannungsquelle zugeführt; seine Impulse werden dann in elektronischen Geräten verstärkt, analysiert und registriert. Für Zählrohre und Szintillationszähler wird eine Gleichspannung zwischen einigen Hundert und einigen Tausend Volt benötigt. Der *Linearverstärker* soll die Detektorimpulse so weit verstärken, daß sie die nachfolgenden elektronischen Geräte steuern können. Dabei ist die Größe der Eingangsimpulse für die einzelnen Detektortypen recht unterschiedlich.

Besonders rauscharme Verstärker sind für Halbleiterdetektoren und Ionisationskammern erforderlich. Für die Impulshöhenanalyse ist es wichtig, daß die Verstärkung linear und mit hoher Stabilität erfolgt.

Impulshöhendiskriminatoren werden benutzt, um Impulse in ihrer Größe zu unterscheiden. Der *Schwellwertdiskriminator* gibt einen Normimpuls ab, wenn der Eingangsimpuls eine eingestellte Schwelle überschreitet. Beim *Einkanalanalysator* werden dagegen nur Impulse innerhalb eines festgelegten Impulskanals registriert. Besteht ein linearer Zusammenhang zwischen Impulshöhe und Strahlungsenergie, wie z. B. beim Szintillationszähler, kann damit das Energiespektrum einer Kernstrahlung bestimmt werden. Da nur die Impulse Berücksichtigung finden, die in den gewählten Impulskanal fallen, ist die Meßzeit bei intensitätsschwachen Strahlern groß, und bei kleinen Halbwertszeiten muß der durch den zeitlichen Abfall bedingte Effekt rechnerisch erfaßt werden. Besser sind *Vielkanalanalysatoren*. Der gesamte Impulshöhenbereich wird unterteilt und jeder ankommende Impuls wird entsprechend seiner Höhe in den zugehörigen Kanal einsortiert.

Beim Registrieren der Impulse sind hohe Zählgeschwindigkeiten erforderlich, da die Impulse bei hohen Strahlungsintensitäten sehr schnell aufeinander folgen. Außerdem wird die Strahlung radioaktiver Nuklide in statistischer Folge emittiert, so daß Impulsabstände auftreten können, die wesentlich kleiner sind als der durch die Strahlungsintensität bestimmte mittlere Impulsabstand. Man benutzt daher heute durchweg elektronische *Zählstufen* mit Zählfrequenzen von einigen 10 MHz, die hintereinandergeschaltet ein beliebig großes Zählvolumen erreichen und in Verbindung mit einer elektronischen Uhr die Impulszahl je Zeiteinheit, die *Impulsrate*, anzeigen.

3.4.3.4. Strahlenschutz

3.4.3.4.1. *Strahlendosimetrie.* Da die Einwirkung ionisierender Strahlung auf Organe und Körperteile des Menschen von keiner Sinneswahrnehmung begleitet ist, bezieht man die beobachteten biologischen, medizinischen und chemischen Strahlenwirkungen auf eine physi-

kalische Größe. Geeignet dafür ist die im Gewebe freigesetzte Energie, die Energiedosis, definiert durch $D = \mathrm{d}E_D/\varrho\,\mathrm{d}V = \mathrm{d}E_D/\mathrm{d}m$. ϱ ist die Substanzdichte und $\mathrm{d}E_D$ die im Volumen $\mathrm{d}V$ mit der Masse $\mathrm{d}m$ absorbierte Energie. Die Maßeinheit ist das *Gry* (Gy), 1 Gy = 1 $\mathrm{Jkg^{-1}}$, es löst das bislang übliche *Rad* (rd), 1 rd = 10^{-2} Gy, ab. Die Energiedosis erfaßt nur die insgesamt absorbierte Menge, nicht die Art der Strahlung und die Größe der bei *einem* Elementarprozeß übertragenen Energie. Die biologischen Wirkungen hängen jedoch auch von der zeitlichen Verteilung der Energieabsorption und vor allem von der Strahlungsart ab. Für den Strahlenschutz wurden daher für die verschiedenen Strahlungsarten dimensionslose *Qualitätsfaktoren* QF eingeführt und die Äquivalentdosis D_{QF} durch das Produkt $QF \cdot D$ definiert. Sie wird in der besonderen Einheit rem [*r*oentgen *e*quivalent *m*edical, engl.] angegeben. Für einen Qualitätsfaktor $QF = 20$ entspricht einer Dosis von 10^{-2} Gy eine Äquivalentdosis von 20 rem. Röntgen-, γ- und β-Strahlung haben einen QF von 1, α- und Protonenstrahlung von 10, Neutronen, je nach ihrer Energie, von 2 bis 10 und schwere geladene Teilchen hoher Energie von 10 bis 20.

Als *Dosisleistung* wird der aus den bislang eingeführten Dosen gebildete Momentanwert $\mathrm{d}D/\mathrm{d}t$ bezeichnet.

Da im allgemeinen die Energiedosis im Gewebe sich nicht direkt messen läßt, führt man in der Regel die Messung in einem Ersatzmedium durch, z. B. durch Bestimmung der *Ionendosis* in Luft und *Ionisationsäquivalente* für die einzelnen Gewebearten.

3.4.3.4.2. *Strahlenschutzverordnung.* Sie dient der Organisation des Schutzes der Bevölkerung vor der schädigenden Einwirkung ionisierender Strahlen und von Personen, die beruflich einer Strahlenbelastung ausgesetzt sind, sowie dem Schutz der Umwelt und von Sachgütern vor radioaktiver Verunreinigung.

Man unterscheidet verschiedene Arten der Strahlenbelastung. *Äußere* oder *innere Strahlenbelastung*, je nachdem ob eine örtliche Trennung zwischen Strahlungsquelle und Organismus besteht oder nicht, *Ganz-* oder *Teilkörperbelastung*, *Kurzzeit-* oder *chronische Strahlenbelastung* u. a. Die Strahlungsquellen werden in geschlossene und offene unterteilt. *Geschlossene Strahlungsquellen* sind so in eine Umhüllung eingeschlossen, daß unter voraussehbaren Betriebs- und Abnutzungsbedingungen eine Verbreitung der radioaktiven Stoffe in die Umgebung ausgeschlossen ist; alle anderen radioaktiven Stoffe sind *offene Strahlungsquellen.* Die gesetzlichen Anordnungen schreiben eine ständige Kontrolle und Überwachung der Labors, der Arbeitsplätze und vor allem der beruflich strahlenexponierten Personen durch staatliche Stellen vor. Bei der Bevölkerung werden drei *Kategorien* und bei strahlenbelasteten Organen und Körperteilen vier *Gruppen* unterschieden. Zur *Kategorie A* zählen Personen, die während ihrer beruflichen Tätigkeit in einem *Kontrollbereich* strahlenexponiert sind, zur *Kategorie B*, die in einem *Überwachungsbereich* tätig sind und zur *Kategorie C* schließlich alle übrigen Personen aus der Bevölkerung. Für die Einteilung in Gruppen gilt: *Gruppe I*: Gesamtkörper, Gonaden und blutbildendes System; *Gruppe II*: Augenlinsen, Magen-Darm-Kanal, Leber, Milz, Nieren, Lunge, Muskeln, Fettgewebe und andere Organe des Körpers, ausgenommen die der Gruppen I und III; *Gruppe III*: Knochen, Schilddrüse und Haut; *Gruppe IV*: Hände, Unterarme und Füße. Kontrollbereiche sind abzugrenzen und zu kennzeichnen, wenn die Strahlenbelastung des Gesamtkörpers einer Person mehr als 0,5 rem $\mathrm{a^{-1}}$ betragen kann; in Überwachungsbereichen bis zu 0,5 rem $\mathrm{a^{-1}}$. In beiden Bereichen dürfen sich Personen aus der Bevölkerung nicht unkontrolliert aufhalten.

Für die verschiedenen Kategorien und Gruppen sind maximal zulässige Äquivalentdosen (MZD) für ionisierende Strahlung (Tab. 3.4.3-8) sowie maximal zulässige Konzentrationen (MZK) von Radionukliden in Luft, Wasser, Rohstoffen, Halbfabrikaten und Erzeugnissen festgelegt.

Tab. 3.4.3-8 Maximal zulässige Dosisäquivalente MZD in rem

Kategorie der Strahlenbelastung	I. Gruppe in Monaten		II. Gruppe in Monaten		III. Gruppe in Monaten		IV. Gruppe in Monaten	
	3	12	3	12	3	12	3	12
A	3	5	8	15	15	30	40	75
B	–	0,5	–	1,5	–	3	–	7,5
C	–	0,5	–	1,5	–	3	–	–

3.4.3.4.3. *Möglichkeiten des Strahlenschutzes.* Bei Arbeiten mit radioaktiven Präparaten gilt die Grundregel: „*Auf keinen Fall in direkten Kontakt kommen*". Radioaktive Substanzen müssen daher mit Zangen, Pinzetten oder speziellen Manipulatoren gehandhabt werden. Da die Dosisleistung mit dem Quadrat der Entfernung von der Strahlenquelle abnimmt, ist der Abstand vom Präparat der einfachste und sicherste Schutz vor Strahlenbelastung. Kann ein Sicherheitsabstand nicht eingehalten werden, so muß die Strahlungsquelle durch Beton- oder Bleiwände abgeschirmt werden. Wegen ihrer großen Ionisationsdichte lassen sich α- und β-Strahlen relativ leicht abschirmen; bereits eine

15 mm dicke Plexiglasscheibe genügt, um auch die energiereichsten zu absorbieren; allerdings kann dabei eine sekundäre Bremsstrahlung auftreten. Für γ-Strahlung sind die Schutzmaßnahmen aufwendiger; sie müssen je nach Nuklid und Aktivität festgelegt werden. Dosisleistung und Aufenthaltszeit sind so aufeinander einzustellen, daß die maximal zulässigen Dosisäquivalente nicht überschritten werden.

3.4.3.5. Herstellung von Radionukliden

3.4.3.5.1. *Aktivierung.* Überwiegend werden Radionuklide durch (n,γ)-Reaktionen mit *thermischen* sowie durch (n,p)- und (n,α)-Reaktionen mit *schnellen Neutronen* im *Reaktor* hergestellt. Darüber hinaus entstehen bei der *Kernspaltung* (n,f) im Reaktor eine große Anzahl *Spaltprodukte*, aus denen ebenfalls technisch wichtige Radionuklide gewonnen werden. Andere Radionuklide werden durch Beschuß mit geladenen Teilchen, z. B. im Zyklotron, erzeugt (Tab. 3.4.3-9).

Als Ausgangsmaterial dienen die natürlichen Elemente oder ihre Verbindungen, in Sonderfällen auch angereicherte Isotope. Nach der Aktivierung werden in der Regel die gewünschten Radionuklide chemisch vom Ausgangsmaterial abgetrennt.

Die durch (n,γ)-Reaktionen gebildeten Radionuklide sind mit dem Ausgangsmaterial isotop, so daß in den meisten Fällen die spezifische Aktivität gering ist; beispielsweise wird bei der ^{31}P(n,γ)^{32}P-Reaktion nur ein kleiner Anteil von P-31 in radioaktiven Phosphor P-32 umgewandelt. Benutzt man dagegen die Reaktion ^{32}S(n,p)^{32}P mit schnellen Reaktorneutronen, dann wird zwar auch nur wenig Radiophosphor gebildet, es ist aber kein anderes Phosphorisotop vorhanden, und der radioaktive Phosphor läßt sich vom Schwefel chemisch abtrennen, so daß seine spezifische Aktivität groß ist. Hohe spezifische Aktivitäten werden für medizinische und biologische Anwendungen gebraucht, da oft aus toxischen Gründen nur geringste Mengen verwendet werden dürfen. Derartige Präparate müssen trägerfrei oder trägerarm sein. Sie lassen sich durch Kernspaltung, durch Reaktionen mit geladenen Teilchen wie p, d, t, α, durch (n,p)- und (n,α)-Reaktionen, durch Reaktionsketten sowie durch den Szilard-Chalmers-Effekt gewinnen. Beim SC-Effekt werden die Atome nach dem Neutroneneinfang aus ihrer Molekülbindung herausgeschlagen und so chemisch isoliert; z. B. entsteht durch Bestrahlung von Iodethyl das Radionuklid I-128, das als freies Atom durch Extraktion leicht von Iodethyl abzutrennen ist.

Tab. 3.4.3-9 Radionuklide *RN* für radiometrische und radiographische Anwendungen: *S* Strahlungsart, $t_{1/2}$ Halbwertszeit, E_s Strahlungsenergie in MeV, $d_{1/2}$ Halbwertsdicke in (mgcm^{-2}); *) maximale Werte

RN	*S*	$t_{1/2}$	E_s	$d_{1/2}$
C-14	β	5570 a	0,156	3,00 (Al)
S-35	β	87 d	0,167	3,0
Ca-45	β	165 d	0,26	6,4
Pm-147	β	2,6 a	0,22	5,3
Kr-87	$\beta(\gamma)$	10,6 a	0,67	27
Tl-204	β	4,1 a	0,77	30
Sr-90 + Y-90	β	28 a	2,2	110
Ce-144 + Pr-144	$\beta(\gamma)$	285 d	3,0	150
Ru-106 + Rh-106	$\beta(\gamma)$	1,0 a	3,5	170
Radium	γ (β, α)	1601 a	2,4	21000 (Fe)
Tm-170	$\gamma(K)$	127 d	0,08*)	1000
Se-75	$\gamma(\beta)$	121 d	0,40*)	9000
Ir-192	$\gamma(\beta, K)$	74 d	0,61*)	12000
Cs-137	$\gamma(\beta)$	30 a	0,66*)	13000
Co-60	$\gamma(\beta)$	5,26 a	1,33*)	19000

3.4.3.5.2. *Radiosynthese.* Spezifisch für sie ist, daß in einer bestimmten Position der betreffenden Molekülstruktur das inaktive Atom durch ein radioaktives Isotop substituiert wird. *Radioisomere Moleküle* unterscheiden sich dadurch, daß das Radionuklid in verschiedenen Positionen auftritt, z. B. gibt es von der Essigsäure CH_3COOH bei der Markierung mit Radiokohlenstoff C-14 vier Radioisomere, je nachdem, ob kein, das eine oder das andere C-Atom oder beide C-Atome durch C-14 ersetzt sind. Radiosynthesen werden meist im Mikrobereich durchgeführt, um wirtschaftlich hohe spezifische Aktivitäten zu erreichen.

3.4.3.6. Radionuklide als Indikatoren

3.4.3.6.1. *Indikatormethode.* Bei dieser 1913 von Hevesy und Paneth entwickelten Methode werden Radionuklide zur Markierung einer Substanz, eines Versuchsobjekts, verwendet. In der einfachsten Form wird ein geschlossenes radioaktives Präparat an einer Probe befestigt oder an eine bestimmte Stelle gebracht. Es kann z. B. der Verschleiß der Auskleidung von Hochöfen kontrolliert werden, wenn in die Auskleidung langlebige Radionuklide eingebaut sind, die bei einem bestimmten Verschleißgrad vom Eisen entfernt werden. Aus der Strahlenmessung im ausfließenden Eisen bzw. in der Auskleidung läßt sich dieser Zeitpunkt genau bestimmen. Für die Reinigung von Pipelines werden sogenannte Molche benutzt, die sich mitunter an besonders stark verschmutzten Stellen festsetzen. Durch Markierung dieser Molche mit Co-60 in Form kleiner Cobaltstäbchen kann ihre Ortung erleichtert werden. Zur chemischen Markierung genügt in den meisten Fällen das Mischen der radioaktiven mit den stabilen Isotopen eines

Elements; nur in speziellen Fällen ist eine Radiosynthese erforderlich. Bei dieser Markierungsform wird das chemisch gleiche Verhalten der Isotope eines Elements ausgenutzt. Nach BRODA und SCHÖNFELD wird die erste Form der Markierung als analytische Indikatormethode bezeichnet, da die Auswahl der verwendeten Radionuklide nur nach physikalisch-technischen Gesichtspunkten erfolgt und lediglich der empfindliche Nachweis des Radionuklids ausgenutzt wird. Die zweite, die kinetische Form der Indikatormethode, benutzt dagegen eben die Möglichkeit, chemisch gleiche Atome, d. h. Isotope, zu unterscheiden. Damit lassen sich z. B. Austauschvorgänge quantitativ an Systemen untersuchen, die hinsichtlich ihrer chemischen Zusammensetzung stationär sind.
Vorteile der Indikatormethode sind der bequeme, schnelle und meist zerstörungsfreie Nachweis der Radionuklide im markierten Objekt bzw. System, die außerordentliche Nachweisempfindlichkeit sowie die Möglichkeit, zwischen Atomen eines Elements zu unterscheiden.

3.4.3.6.2. *Emaniermethode.* Sie beruht auf der natürlichen Radioaktivität der Isotope Rn-222 und Rn-220 des Edelgases Radon, das auch Emanation genannt wird. Bei der Herstellung der Untersuchungsproben wird ein Radionuklid, das dieses Edelgas abgibt, z. B. Radium oder Thorium, möglichst homogen mit der zu untersuchenden Substanz vermischt. Die gebildete Emanation entweicht sowohl durch Diffusion entlang der Korngrenzen (Oberflächendiffusion) und durch die Körner (Gitterdiffusion) als auch durch die Rückstoßenergie beim α-Zerfall aus der Probe. Aus dem gemessenen Verhältnis der ausgetretenen Emanation zur insgesamt gebildeten (Emaniervermögen) gewinnt man Aussagen über die Struktur der Probe; so ist z. B. das Emaniervermögen bei oberflächenreichen und amorphen Substanzen größer als bei oberflächenarmen und kristallinen Substanzen. Eine Variante dieser Methode besteht darin, daß man ein radioaktives Edelgas, z. B. Rn oder Kr-85, durch eine Gasentladung in oberflächennahe Schichten des Festkörpers einbaut und die Abgabe unter verschiedenen Bedingungen untersucht. Diese Methode, zuerst von HAHN beschrieben, ist außerordentlich empfindlich und experimentell leicht zu realisieren.

3.4.3.6.3. *Radioreagenzmethode.* Bei ihr reagieren die zu bestimmenden oder zu prüfenden Substanzen mit radioaktiv markierten Reagenzien, und das Reaktionsprodukt wird über seine Aktivität bestimmt. Durch den Einsatz radioaktiver Reagenzien kann z. B. die Empfindlichkeit der Titration und der Papierchromatographie erheblich gesteigert werden; so daß noch Substanzen nachgewiesen werden können, die weit unter der Erkennbarkeit durch Farbreaktionen liegen. Diese Methode, die ursprünglich von EHRENBERG auf der Basis von Radioblei ausgearbeitet wurde, erfuhr mit der Entdeckung der künstlichen Radioaktivität eine wesentliche Erweiterung und ist heute ein fester Bestandteil der analytischen Chemie.

3.4.3.6.4. *Isotopenverdünnungsmethode.* Bei der *einfachen* oder *direkten* Methode wird dem zu analysierenden Stoffgemisch eine bekannte Menge des zu bestimmenden Stoffs in markierter Form mit bekannter spezifischer Aktivität zugesetzt. Nachdem sich das Mischungsgleichgewicht eingestellt hat, wird ein Teil des Stoffs nach den üblichen chemischen Trennverfahren, oft in mehreren Stufen, bis zur sicheren Reinheit isoliert. Aus den bekannten bzw. gemessenen spezifischen Aktivitäten des zugesetzten und des isolierten Stoffs sowie aus den Massenverhältnissen kann dann die unbekannte Stoffmenge bestimmt werden. Ist m_x die zu bestimmende Menge im Stoffgemisch, m_0 die Menge der markierten Zugabe und A_{s0} ihre spezifische Aktivität sowie m_1 die Menge des abgetrennten reinen Stoffs und A_{s1} ihre spezifische Aktivität, so gilt die Beziehung $(m_x + m_0)/m_1 = m_0 A_{s0}/(m_1 A_{s1})$, d. h. $m_x = m_0[(A_{s0}/A_{s1}) - 1]$; ist $A_{s0} \gg A_{s1}$, dann gilt $m_x = m_1 A_0/A_1$, und das Aktivitätsverhältnis gibt dann direkt die Ausbeute des Trennverfahrens an. Mit dieser Methode können Stoffe in Systemen quantitativ bestimmt werden, ohne sie vollständig zu isolieren. Sie ist daher besonders geeignet für die Analyse chemisch ähnlicher Substanzen, z. B. biologischen, bei denen die herkömmlichen Bestimmungsmethoden nur mit großem Aufwand zum Ziel führen. Die *inverse Isotopenverdünnungsmethode* wird zur Analyse sehr geringer Stoffmengen und für die Bestimmung von radioaktiven Substanzen in Gemischen aktiver Stoffe eingesetzt. Im ersten Fall muß der zu bestimmende Stoff vorher radioaktiv markiert werden (↑ 3.4.3.5.1.). Die Abtrennung erfolgt dann durch Zugabe einer bekannten Menge des zu bestimmenden Stoffs in inaktiver Form.

3.4.3.6.5. *Aktivierungsanalyse.* Sie dient zur qualitativen und quantitativen Analyse von Substanzen mit Hilfe der in ihnen durch Aktivierung erzeugten künstlichen Radioaktivität. Die Aktivierung kann mit Neutronen, geladenen Teilchen oder γ-Strahlung erfolgen; die Neutronenaktivierungsanalyse mit einem Kernreaktor oder Neutronengenerator ist dabei am gebräuchlichsten. Die erzeugten Radionuklide sind qualitativ durch Art, Energie und Halbwertszeit der emittierten Strahlung sowie quantitativ durch die Aktivität mit hoher Selektivität und Empfindlichkeit nachweisbar. Bei einem Neutronenfluß Φ und der Bestrahlungsdauer t_A bildet sich bei N Kernen eines bestimmten Iso-

tops mit einem Aktivierungsquerschnitt σ die Aktivität $A = \sigma \Phi N[1 - \exp(-\ln 2 \cdot t_A/t_{1/2})]$. $t_{1/2}$ ist die Halbwertszeit des gebildeten Radionuklids; sind σ, Φ, $t_{1/2}$ und t_A bekannt, läßt sich daraus die Anzahl N der gesuchten Atome bestimmen.

Da jedoch die Bestrahlungsparameter und die kernphysikalischen Daten nicht stets mit genügender Sicherheit verfügbar sind, wird vorzugsweise die Relativmethode benutzt. Man bestrahlt eine Standardprobe mit bekannter Substanzmenge m_S unter möglichst gleichen Bedingungen wie die Analysenprobe. Die unbekannte Substanzmenge m_P kann dann durch die gemessene Aktivität von Probe A_P und des Standards A_S, nachdem diese auf gleiche Meßbedingungen (Abklingzeit, chemische Ausbeute) normiert wurden, nach $m_P = m_S A_P / A_S$ ermittelt werden.

Eine Variante der Relativmethode ist die des „inneren" Standards. Hierbei wird vor der Aktivierung die zu analysierende Probe mit der als Standard gewählten Substanz bekannter Menge homogen gemischt. Als idealer innerer Standard gilt das zu bestimmende Element selbst, weil dann Standard und Probe wirklich gleichen Aktivierungsbedingungen unterliegen. Dazu teilt man die Analysenprobe in mindestens zwei Teile und bestrahlt einen Teil direkt und den anderen nach Zugabe einer definierten Menge des zu bestimmenden Elements. Die unbekannte Substanzmenge ergibt sich dabei zu $m_{Pr} = m_{St}/[(A_{PS}/A_{Pr}) - 1]$, wobei die Aktivitäten A_{Pr} und A_{PS} den Proben „ohne" und „mit" innerem Standard zuzuordnen sind; m_{St} ist die zugegebene Menge.

In der Ausführung unterscheidet man die chemische von der zerstörungsfreien Aktivierungsanalyse. In der *zerstörungsfreien Aktivierungsanalyse* wird die Aktivität des für die Bestimmung einer Substanz in der Probe wichtigen Radionuklids direkt am bestrahlten Material gemessen. Nach der Energie mit Hilfe einer energieselektiven Methode der γ- oder β-Spektrometrie und nach der Halbwertszeit der entsprechenden Strahlung – meist γ-Strahlung – lassen sich das Radionuklid und damit die Substanz identifizieren. Voraussetzung dazu ist, daß eine eventuelle Überlagerung von Strahlungskomponenten verschiedener Radionuklide meßtechnisch zu trennen ist. Vorteilhaft sind dabei der geringe Zeitaufwand und die Automationsfähigkeit. Aktivierungsanalysen am Neutronengenerator werden durchweg nach dieser Methode durchgeführt. Die kurzen Halbwertszeiten der entstehenden Radionuklide erlauben dabei keine zeitraubenden Operationen, und die Störung durch Fremdaktivität ist oft sehr gering.

In der *chemischen Aktivierungsanalyse* werden die interessierenden Radionuklide aus der bestrahlten Probe auf chemischem Wege abgetrennt, ehe ihre Aktivität bestimmt wird. Eine Trennung bis zur radiochemischen Reinheit dominierte bis etwa 1955; seitdem ist durch den zunehmenden Einsatz von γ-Spektrometern die Kombination beider Methoden in den Vordergrund getreten. Durch einfache Gruppentrennung kann ein großer Teil der nicht interessierenden Störaktivitäten beseitigt werden, so daß danach eine spektrometrische Analyse relativ einfach möglich ist.

Da die Substanzmengen oft äußerst gering sind, setzt man vor Beginn der chemischen Operation dem aufzuarbeitenden Substanzgemisch inaktive Träger zu, die mit den radioaktiven Komponenten chemisch identisch sind, um Absorptions- und Löslichkeitsverluste zu vermeiden. Trägerzusatz und radioaktive Komponente müssen vor der Trennung gleiche chemische Form haben. Die Trennung selbst braucht bei der Trägertechnik nicht quantitativ zu erfolgen, da am Ende der Analyse die Ausbeute bestimmt werden kann. Die chemische Aufarbeitung erfolgt meist durch Fällung, Extraktion, Ionenaustausch, Destillation, Chromatographie oder Elektrolyse. Der größte Nachteil der chemischen Aktivierungsanalyse ist der relativ große Zeitaufwand. Erst durch die Entwicklung spezieller chemischer Schnellmethoden wurden auch einige kurzlebige Radionuklide der Analyse zugänglich. Ihr Hauptvorteil liegt in der anders unerreichbaren Empfindlichkeit als Folge der fehlenden Störaktivitäten.

Die Aktivierungsanalyse ist heute eines der wichtigsten Verfahren der modernen Analytik und wird in nahezu allen Industriezweigen und Forschungsstätten angewendet; auch zur Produktionskontrolle und vor allem zur Bestimmung von Spurengehalten in Reinststoffen wird sie eingesetzt, z. B. von Metallen, Legierungen oder Reagenzien in biologischen, organischen und geologischen Stoffen sowie in Nahrungsmitteln. Für Spurengehalte unter 10^{-3} Gewichtsprozent ist oft nur sie geeignet; aber auch bei Gehalten von $10^{-3}\%$ bis 1% wird sie oft den üblichen Analysenmethoden vorgezogen, da sie bei gleicher Genauigkeit schneller und automationsfähig ist.

3.4.3.6.6. *Methode der induzierten Kernreaktion.* Hier werden die bei einer Kernreaktion spontan emittierten Teilchen oder γ-Quanten für eine Analyse genutzt. Man bestimmt aus ihnen die Kernreaktion und das an dieser beteiligte Element; z. B. kann bei (n,γ)-Reaktionen, wo der Produktkern nicht radioaktiv ist, die entstehende γ-Strahlung zur Analyse benutzt werden. Mit den Methoden der *Neutronen-Gamma-Karottage* werden Elemente im Bohrloch durch die γ-Strahlung identifiziert, die sie beim Einfang von Neutronen emittieren. Der Nachweis von Bor kann über die Reaktion $^{10}B(n,\alpha)^7Li$ durch Messung der entstandenen

α-Strahlung erfolgen. Die Methode der induzierten Kernreaktionen stellt eine wertvolle Ergänzung zur aktivierungsanalytischen Methode dar.

3.4.3.7. Radionuklide als Strahlenquellen

3.4.3.7.1. *Radioaktive Strahlenquellen.* Bei der Anwendung radioaktiver Nuklide als Strahlenquellen sind Art und Energie der Strahlung sowie die Halbwertszeit die bestimmenden Parameter. Das Element, zu dem das Radionuklid gehört, ist hierbei unwesentlich. Die Wahl des Radionuklids und die technische Form der Strahlenquelle werden durch den Charakter des Meßobjekts, seine Abmessungen, die geforderte Genauigkeit und Schnelligkeit der Messung beeinflußt. Im allgemeinen werden geschlossene Strahlenquellen eingesetzt, wobei die Halbwertszeit des Radionuklids mehrere Jahre beträgt und nur in Ausnahmefällen kleiner als ein Jahr ist (↑ Tab. 3.4.3-9).

3.4.3.7.2. *Radiometrie.* Grundlage der Radiometrie ist die *Absorption* und *Streuung* der Kernstrahlung durch die Atome des Stoffs, den sie durchdringt; dabei gehen Art und Energie der Strahlung, Materialart und Schichtdicke als Parameter ein. Die Meßverfahren zur Bestimmung der Dicke, der Dichte und anderer Größen beruhen danach auf der *Durchstrahlung* oder auf der *Rückstreuung*, α-Teilchen werden aufgrund ihres großen spezifischen Ionisationsvermögens bereits in sehr dünnen Schichten vollständig abgebremst und absorbiert; sie eignen sich zur Dickenbestimmung von extrem dünnen Folien nach dem Durchstrahlungsverfahren. Da β-Teilchen durchdringungsfähiger als α-Teilchen sind, sind sie deshalb besonders geeignet für radiometrische Untersuchungen an dünnen Schichten sowohl nach dem Durchstrahlungs- als auch nach dem Rückstreuverfahren, mit dem sich auch Dickenbestimmungen von Schichten, die auf Unterlagen aufgetragen sind, durchführen lassen. Dagegen werden γ-Strahlen vorwiegend zur Bestimmung großer Materialdicken und -dichten eingesetzt.

3.4.3.7.3. *Radiographie.* Nach dem Absorptionsgesetz werden γ- oder Röntgenstrahlen je nach Dicke und Dichte eines durchstrahlten Gegenstands geschwächt, so daß der Schwärzungsgrad eines *Radiogramms* auf einer fotografischen Emulsion Materialfehler erkennen läßt; eine zerstörungsfreie Materialprüfung ist dadurch also möglich. Um eine gute *Abbildungsschärfe* zu erzielen, müssen Objekt und Film engen Kontakt, Strahlenquelle und Objekt jedoch einen möglichst großen Abstand voneinander haben. Punktförmige Strahlenquellen sind deshalb zu bevorzugen und die spezifische Aktivität der eingesetzten Radionuklide muß möglichst groß sein. Für die *Registrierung* werden schnelle doppelseitige Emulsionen und Röntgenfilme benutzt, die meist noch in Verstärkerfolien aus Blei eingeschlossen sind. Die *gebräuchlichsten Strahlenquellen* sind Co-60, Cs-137, Ir-192 und neuerdings auch Am-241 und Eu-155. Eine spezielle Methode der Radiographie ist die *Autoradiographie* (Tafel 10). Bei ihr wird die Probe, die entweder selbst radioaktiv ist oder radioaktiv markiert worden ist, in engen Kontakt mit einer fotografischen Emulsion gebracht. Das erhaltene *Autoradiogramm* stellt die Verteilung des Radionuklids in der Probe dar.

3.4.3.7.4. *Strahlenschranken.* Eine vereinfachte Form der Dichtemessung wird in der Regel- und Steuertechnik als Strahlenschranke eingesetzt. Damit kann man nicht nur den Füllstand in Behältern mit festem oder flüssigem Füllgut messen und kontrollieren, sondern auch Trennschichten verschiedener Stoffe ermitteln, Bandförderanlagen oder Durch- und Einfahrten überwachen, Stücke zählen u. a. Die Wahl der Strahlenquelle richtet sich nach dem zu durchstrahlenden Stoff; meist kommen nur γ-Strahler in Frage, da neben dem Meßgut oft Wände aus Stahl oder ähnlichem Material durchstrahlt werden müssen. Als Strahlenquelle werden daher bevorzugt Co-60 und Cs-137 und als Detektor das Geiger-Müller-Zählrohr benutzt.

3.4.3.7.5. *Strahlenwirkung.* Die Wechselwirkung der radioaktiven Strahlung mit Materie äußert sich nicht nur in einer Änderung der Energie, der Intensität und der Richtung der Strahlen, sondern auch in Änderungen der Materie selbst. Durch die aufgenommenen beträchtlichen Energien können sich die chemischen und physikalischen Eigenschaften des Stoffs ändern; z. B. durch Anregung oder Ionisation der Atome des Stoffs bei Teilchenstrahlung mit Ladung. Schnelle Neutronen dagegen rufen durch Stöße Gitterstörungen sowie Kernreaktionen hervor; γ-Strahlen zeigen nur eine indirekte Strahlenwirkung durch die erzeugten energiereichen Elektronen nach dem Foto-, dem Compton- und dem Paarbildungseffekt.

Metalle lassen sich praktisch nur durch Neutronen beeinflussen; dabei ändern sich z. B. der elektrische Widerstand und der Elastizitätsmodul. Halbleiterkristalle können durch die erzeugten Gitterstörungen sogar ihren Leitungscharakter ändern. Starke Änderungen treten bei Hochpolymeren auf; vor allem Spaltung oder Vernetzung der Makromoleküle. Polyethylen, Polystyrol, Dederon, Kautschuk z. B. neigen zur Vernetzung; sie erhalten bessere thermische und mechanische Eigenschaften.

Bei Plexiglas, Teflon u. a. dagegen tritt Spaltung der Vernetzung auf. In lebenden Organismen führt die Bestrahlung mit sehr großen Dosen zur Sterilität bzw. zum Tod. Die *letale Dosis* schwankt zwischen 6 Gy für Menschen und $2 \cdot 10^4$ Gy für sporenbildende Bakterien. Nahrungs- und Arzneimittel sowie chirurgische In-

strumente können daher durch Strahlung sterilisiert werden. In der Landwirtschaft kann Bestrahlung von Saatgut zur Ertragssteigerung und zur Züchtung neuer Sorten dienen. Hauptziel der *Strahlentherapie* ist es, kranke Zellen in Krankheitsherden, in Organen oder in Gewebeteilen, zu zerstören. Eine Schädigung des gesunden Gewebes sucht man durch eine alternierende Bestrahlung mit kleinen Dosen über längere oder durch einen zyklischen Wechsel der Bestrahlungsrichtung zu vermeiden. In ihrer technischen Ausführung sind diese Bestrahlungsanlagen als Cobalt- bzw. Caesiumkanone bekannt geworden. In Gasen erzeugt ionisierende Kernstrahlung Ionen, vergrößert so deren Leitfähigkeitsvermögen und beseitigt elektrostatische Aufladungen, die durch Reibung in isolierenden Stoffen entstehen; z. B. in der Papier-, Kunststoff- und Gummiindustrie sowie in der pulververarbeitenden Industrie. Radioaktive Lichtquellen und Batterien beruhen ebenfalls auf der ionisierenden Wirkung der radioaktiven Strahlung.

3.4.3.8. Altersbestimmungen durch Radionuklide

3.4.3.8.1. *Radiokohlenstoffmethode.* Durch Wechselwirkung der kosmischen Strahlung mit der oberen Atmosphäre entstehen schnelle Neutronen, die aus dem Luftstickstoff N-14 radioaktiven Kohlenstoff C-14 erzeugen. Als Kohlendioxid wird dieser von Wind und Meeresströmungen gleichmäßig verteilt, so daß sich durch Radiokohlenstoffassimilierung im Laufe der Erdgeschichte ein konstanter C-14-Gehalt in allen lebenden organischen Stoffen eingestellt hat. Mit dem Tod hört dann die Assimilation auf, und der radioaktive Zerfall des C-14 wird nicht mehr ausgeglichen. So ist die spezifische Aktivität eines vor 5570 Jahren gestorbenen Lebewesens nur noch halb so groß wie die eines lebenden Organismus, weil die Halbwertszeit von C-14 5570 Jahre beträgt. Die ältesten Funde, die mit dieser Methode zuverlässig auf 4900 Jahre datiert wurden, stammen aus der ersten ägyptischen Dynastie, aus der Zeit HEMAKAS und DJETS.

3.4.3.8.2. *Tritiummethode.* Ebenso wie C-14 wird auch Tritium in der Atmosphäre durch Einwirkung der Höhenstrahlung erzeugt. Mit den Niederschlägen gelangt es in die Gewässer und bildet hier durch den ständigen Austausch mit der Atmosphäre ebenfalls einen konstanten Tritiumgehalt. In Wässern dagegen, die ihren Kontakt mit der Atmosphäre verloren haben, wird der Tritiumgehalt mit der Halbwertszeit von 12,3 Jahren abnehmen. Auf diese Weise ist eine Altersbestimmung von eingesickertem Wasser, von Quellwasser oder Grundwasser möglich. Auch das Alter von Weinen kann auf diese Weise ermittelt werden. Wegen der relativ geringen Halbwertszeit sind jedoch nur Alter bis zu etwa 50 Jahren bestimmbar.

3.4.3.8.3. *Bleimethode.* Die bereits vor einigen Milliarden Jahren abgeschlossene Bildung bzw. Umwandlung der chemischen Elemente und ihrer Isotope wird in einigen Fällen durch radioaktive Zerfallsprozesse fortgeführt, so daß hier die Isotopenzusammensetzung von der Zeit abhängt. Die drei natürlichen radioaktiven Zerfallsreihen des *Thoriums*, *Actiniums* und *Uraniums* enden mit den stabilen Isotopen Pb-206, Pb-207, Pb-208 und He-4. Bei der Altersbestimmung nach der Bleimethode wird das Verhältnis der Anteile des radioaktiven Ausgangsnuklids und seines Zerfallsprodukts bestimmt und daraus über die Halbwertszeit des Radionuklids das Alter berechnet.

3.4.3.8.4. *Weitere Möglichkeiten.* Die Altersbestimmung nach der *Heliummethode* geht vom Uranium-, Thorium- und Heliumgehalt der Mineralien aus und beruht auf dem gleichen Prinzip wie die Bleimethode; sie ist jedoch mit Vorbehalt anzuwenden, da praktisch immer mit einem gewissen Verlust an He-4 im Mineral im Laufe geologischer Zeiten zu rechnen ist.

Neben den bereits genannten gibt es noch einige andere natürliche Radionuklide, die sich mit einer großen Halbwertszeit in ein stabiles Nuklid umwandeln und deshalb innerhalb geologischer Zeiten die relative Häufigkeit ihrer stabilen Isotope verändern, z. B. die Radionuklide K-40 und Rb-87. So zerfällt das Rb-87 durch β-Zerfall in das stabile Sr-87, wobei die Halbwertszeit $6 \cdot 10^{10}$ Jahre beträgt. Nach dieser Methode wurde das Alter der ältesten Mineralien auf rund $4 \cdot 10^9$ Jahre datiert. Für die Geologen liefern diese Methoden die einzige absolute Zeitskala und damit Anhaltspunkte über das Alter der Erde.

3.4.4. Neutronenphysik

3.4.4.1. Freies Neutron. Das Neutron gehört als Bestandteil der Kernmaterie zusammen mit dem Proton zur Gruppe der Nukleonen. Seine Masse ist um etwa 2,5 Elektronenmassen schwerer als die des Protons. Obwohl es keine Ladung trägt, besitzt es ein zum Spin des Neutrons proportionales magnetisches Moment μ_n. Das Neutron ist als freies Elementarteilchen unstabil. Es zerfällt mit einer Halbwertszeit von etwa 10,8 min in ein Proton, ein Elektron und ein Antineutrino $n \to p + e^- + \bar{\nu}$. Die Zerfallsenergie (obere Grenze des Elektronenspektrums) beträgt dabei 0,7826 MeV.

3.4.4.2. Erzeugung von Neutronen. *Freie Neutronen* kommen in der Natur nur als Bestandteile der Höhenstrahlung vor. Für physikalische Experimente und für verschiedene Anwendungen müssen sie unmittelbar erzeugt werden. Das ist nicht nur wegen ihrer Instabilität erforderlich, sondern auch deshalb, weil sich Neutronen nicht in einem Gefäß oder in einer physikalischen Apparatur einschließen lassen (s. a.

3.4.4.7.). Neutronen werden mittels Kernreaktionen erzeugt, und zwar mit Hilfe von Beschleunigern, mit radioaktiven Quellen oder in Kernreaktoren.

Tab 3.4.4-1 Neutronenerzeugung mit Beschleunigern

Reaktion	Q	E_s
exotherme Reaktionen		
H-2 (d, n) He-3	$Q = 3{,}28$ MeV	
H-3 (d, n) He-4	$Q = 17{,}6$ MeV	
Li-7 (d, n) $\rightarrow 2\alpha$ Be-8	$Q \approx 15$ MeV	
Be-9 (d, n) B-10	$Q = 4{,}39$ MeV	
endotherme Reaktionen		
H-3 (p, n) He-3	$Q = 0{,}764$ MeV	$E_s = 1{,}02$ MeV
Li-7 (p, n) Be-7	$Q = 1{,}65$ MeV	$E_s = 1{,}88$ MeV
Be-9 (p, n) B-9	$Q = 1{,}85$ MeV	$E_s = 2{,}06$ MeV

Die Kernreaktionen für die Erzeugung von Neutronen in Beschleunigern sind in Tab. 3.4.4-1 zusammengefaßt. Es entstehen intensive Neutronenströme relativ einheitlicher Energie, welche durch die Energie der beschleunigten Ionen, den Q-Wert der Reaktion und die Austrittsrichtung der erzeugten Neutronen bestimmt ist. Da die beschleunigten Ionen die Coulombbarriere des Targetkerns überwinden müssen, werden als Targetmaterialien fast ausschließlich leichte Elemente wie H-2 (Deuterium), H-3 (Tritium), Li und Be verwendet. Für die Erzeugung hochenergetischer (schneller) Neutronen werden die exothermen (d,n)-Reaktionen, für die Erzeugung niederenergetischer (langsamer) Neutronen endotherme (p,n)-Reaktionen dicht oberhalb der Schwellenergie E_s verwendet.
Sehr intensive Neutronenquellen können unter Ausnutzung des Kernfotoeffekts (γ,n) realisiert werden. Die γ-Strahlung wird zweckmäßigerweise mittels intensiver Elektronenbeschleuniger als Bremsstrahlung im gleichen Target erzeugt. Da die durch die Elektronen erzeugte Bremsstrahlung mit wachsendem Atomgewicht des Targetmaterials an Härte und Intensität zunimmt, der Wirkungsquerschnitt für die (γ,n)-Reaktion ebenfalls mit wachsendem Atomgewicht steigt, werden als Targetmaterial die schwersten Elemente (U, Bi) verwendet. Bei Verwendung von U-235 kann die Neutronenproduktion durch Spaltung zusätzlich um etwa eine Größenordnung erhöht werden. Es werden Neutronenflüsse von $10^{18}\ s^{-1}$ im Impuls (Dauer etwa 1 μs) erreicht. Zu den intensivsten Neutronenquellen zählt die Spallation (↑ 3.4.2.5.7.) von Kernen schwerer Elemente im Teilchenstrom von Hochstrombeschleunigern. Nach einem kanadischen Projekt in Chalk River soll damit eine Neutronenflußdichte von $10^{16}\ cm^{-2}s^{-1}$ verfügbar sein, welcher selbst bei Leistungsreaktoren der Zukunft nicht zu erreichen sein wird. Hohe Neutronenflüsse sind für die Erprobung von Konstruktionsmaterialien für Kern- und Fusionsreaktoren von außerordentlicher Bedeutung.
Leicht transportable kleine Neutronenquellen mit allerdings nur geringer Quellenstärke lassen sich mit Hilfe radioaktiver α- oder γ-Strahler herstellen, deren Strahlung so energiereich ist, daß sie durch Kernreaktionen an leichten Elementen Neutronen auslösen können. Die am häufigsten benutzten Neutronenquellen dieser Art verwenden die Reaktion $^{9}Be(\alpha,n)\ ^{12}C$, wobei die auslösenden α-Teilchen durch Radium, Radon, Polonium oder Plutonium geliefert werden. Das Energiespektrum einer solchen Quelle ist kontinuierlich und reicht bis etwa 10 MeV.
Die wichtigste Neutronenquelle ist die aktive Zone des Kernreaktors. Das Energiespektrum ist kontinuierlich und reicht bis zu 10 MeV. Den Hauptanteil bilden jedoch die „thermischen" Neutronen, welche durch Abbremsung der Spaltneutronen im Moderator des Reaktors entstehen. Ihr Geschwindigkeitsspektrum entspricht etwa der thermischen Bewegung der Neutronen im Temperaturgleichgewicht mit den Reaktormaterialien. Der Anteil der *thermischen Neutronen* am Gesamtneutronenfluß hängt vom Reaktortyp ab. *Schnelle Neutronen* werden diejenigen Neutronen genannt, welche direkt aus dem Prozeß der Kernspaltung hervorgehen und noch nicht abgebremst wurden (1 bis 10 MeV). Dagegen sind die *Resonanzneutronen* erst teilweise abgebremste Neutronen mit Energien zwischen 1 eV und 1 MeV.
Die mit Forschungsreaktoren erreichbaren Neutronenflußdichten liegen je nach Leistung und Größe des Reaktors im Inneren der aktiven Zone bei 10^{13} bis $10^{15}\ cm^{-2}s^{-1}$. Durch spezielle Öffnungen können Strahlenbündel aus dem Reaktor herausgeführt werden, in denen Stromdichten von 10^{7} bis $10^{10}\ cm^{-2}s^{-1}$ erreicht werden.
Die Neutronenströme an einem Beschleuniger werden im allgemeinen impulsartig erzeugt. Das ist für eine Reihe von Anwendungsfällen unumgänglich. Auch ein Reaktor läßt unter speziellen Bedingungen einen Impulsbetrieb zu. So wird im Impulsreaktor IBR 2 des VIK Dubna die aktive Zone periodisch kurzzeitig in den überkritischen Zustand gebracht. Bei einer Frequenz von 5 Hz und einer Impulslänge von 150 μs werden im Impuls 10^{16} Neutronen erzeugt. Das entspricht im Mittel der Leistung eines Reaktors von 4 MW, im Impuls jedoch von 8000 MW.

3.4.4.3. Durchgang von Neutronen durch Materie. Da die Neutronen beim Durchgang durch Materie nur mit Atomkernen, nicht aber mit den Hüllenelektronen wechselwirken, verhalten sie sich grundsätzlich anders als geladene

Teilchen. So kann ein Neutron auch in kompakter Materie Weglängen in der Größenordnung von mehreren cm ohne Wechselwirkung zurücklegen. Kommt es zu einem Stoß mit einem Kern der Substanz, tritt Absorption des Neutrons im Kern oder Streuung an ihm ein.

3.4.4.3.1. *Streuung und Absorption.* Bei der *elastischen Streuung* erfolgt die Abbremsung der Neutronen dadurch, daß ein Teil der kinetischen Energie an den gestoßenen Kern übergeht. Dieser Prozeß ist besonders effektiv, wenn es sich um ein Material aus leichten Elementen handelt. Die Abbremsung der Neutronen ist beendet, wenn sich das „Neutronengas" mit den Kernen der umgebenden Substanz im thermischen Gleichgewicht befindet. Stark bremsende und dabei die Neutronen wenig absorbierende Substanzen nennt man *Moderatoren* (↑ 3.5.1.).

Schnelle, d. h. hochenergetische Neutronen, können außerdem kinetische Energie durch *unelastische Streuprozesse* verlieren. Da die unelastische Streuung von der Lage anregbarer Kernniveaus abhängt, eignen sich hierfür mittelschwere und schwere Elemente.

Durch die Absorption verschwindet das Neutron für die hier betrachteten Prozesse, wobei die dafür zuständigen Kernreaktionen in diesem Zusammenhang nicht im einzelnen analysiert zu werden brauchen. Die Summe der entsprechenden Wirkungsquerschnitte $\sigma_{tot} = \sigma_{el} + \sigma_{inel} + \sigma_{abs}$ bestimmt die mittlere freie Weglänge (↑ 2.4.2.4.) der Neutronen im Material $l = 1/n\sigma_{tot}$, wobei n die Zahl der Kerne in der Volumeneinheit bedeutet (Tab. 3.4.4-2).

Tab. 3.4.4-2 Wirkungsquerschnitte für einige Nuklide für thermische Neutronen; *) natürliches Isotopengemisch

gute Moderatoren			*gute Absorber*		vorherr-
	σ_{abs} in mb	σ_{tot} in b		σ_{abs} in b	schende Reaktion
H-1	0,33	38	B*)	755	(n, α)
H-2	0,57	15	Cd*)	2450	(n, γ)
C*)	3	4,8	Gd*)	46000	(n, γ)
O*)	< 0,2	4,2			

Schlechter Adsorber Al $\sigma_{tot} = 0{,}23$ b

3.4.4.3.2. *Neutronenbeugung.* Sie tritt beim Durchgang thermischer Neutronen durch Kristalle auf. Für diese Neutronen ist die Materiewellenlänge von der Größenordnung des Gitterabstands und entspricht der Wellenlänge harter Röntgenstrahlung. Neutronen können daher ebenso wie Röntgenstrahlen an Kristallgittern gebeugt werden (Abb. 3.4.4-3), insbesondere ist eine Bragg'sche Reflexion (6.4.2.3.1.) an einer Netzebenenschar möglich. Diese Erscheinung eignet sich zur Herstellung monochromatischer Neutronenbündel, da für jeden Einfallswinkel nur eine bzw. nur wenige Wellenlängen reflektiert werden, sowie für die Bestimmung der Neutronenenergie (↑ 3.4.4.4.).

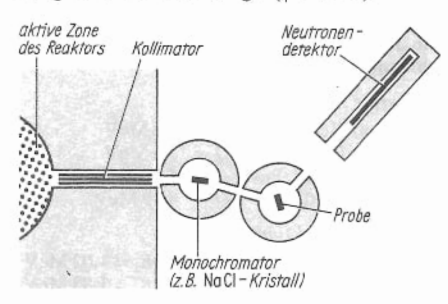


Abb. 3.4.4-3 Prinzipieller Aufbau einer Neutronenbeugungsapparatur; Abschirmungen *gerastert*

Voraussetzung für diese Interferenzerscheinungen ist die Kohärenz (Interferenzfähigkeit) der einzelnen Streuakte. Diese Kohärenz ist jedoch nur zum Teil erfüllt, was durch die Aufteilung des elastischen Streuquerschnitts in einen kohärenten und einen inkohärenten Anteil $\sigma_{el} = \sigma_{koh} + \sigma_{inkoh}$ ausgedrückt wird. Der inkohärente Anteil des Querschnitts hat zwei Ursachen:

1. Bei Kernen des Kristallgitters mit einem Spin $I \neq 0$ treten für ein einfallendes Neutron je nach räumlicher Einstellung dieses Spins verschieden starke Wechselwirkungen ein. Die Verteilung der Einstellung auf die Gitterplätze erfolgt jedoch statistisch. Der kohärente Anteil des Streuquerschnitts von Protonen umfaßt z. B. nur etwa 2% des Gesamtstreuquerschnitts, während er bei Deuteronen etwa 80% beträgt.
2. Bei Kernen des Gitters, welche ein Isotopengemisch bilden, sind auch die einzelnen isotopen Kerne statistisch auf die Gitterplätze verteilt. Da die Streuquerschnitte verschiedener Isotope stark voneinander abweichen können, wird auch hierbei das einfache Bild der Interferenzen bei der Streuung gestört.

Nur isotopenreine Elemente mit dem Kernspin $I = 0$ zeigen reine kohärente elastische Streuung. Einen zusätzlichen Beitrag zum inkohärenten Streuquerschnitt verursachen alle Abweichungen von der idealen Anordnung der Atome im Kristall. Das sind in den real einsetzbaren Einkristallen Versetzungen und andere Gitterstörungen sowie thermische Gitterschwingungen (Phononen).

3.4.4.3.3. *Totalreflexion.* Sie ist eine weitere Besonderheit der Bewegung von Neutronen in kompakter Materie, welche ebenfalls mit ihren Welleneigenschaften verknüpft ist: Neutronen verhalten sich so, als ob dem Material ein Brechungsindex n zugeordnet wäre. Dieser liegt wie

bei der Röntgenstrahlung sehr nahe dem Wert 1 und spielt deshalb im allgemeinen keine Rolle. Für thermische Neutronen gilt im günstigsten Fall (Ni-58) $1 - n \approx 10^{-5}$. Deshalb ist auch die Erscheinung der Brechung von Neutronenwellen an Grenzflächen zwischen verschiedenen Materialien kaum beobachtbar, wohl aber die der Totalreflexion bei streifendem Einfall. Da im allgemeinen $n < 1$ ist, tritt die Totalreflexion für jene Neutronen ein, welche von außen her in das Material eintreten wollen, und zwar für Winkel zur Grenzfläche, welche kleiner als der Grenzwinkel Θ_c sind. Es gilt analog zur Optik $\Theta_c = \sqrt{1 - n^2}$. Der Grenzwinkel liegt für thermische Neutronen bei etwa 20 Bogensekunden.
Der Effekt der Brechung und Totalreflexion ist unabhängig von der Anordnung der Atomkerne im Material und ebenso von Oberflächenrauhigkeiten, sofern die Neutronenwellenlängen die Dimensionen der Rauhigkeiten übersteigen. Er findet in den sogenannten *Neutronenleitern* Anwendung, welche an verschiedenen Forschungsreaktoren installiert sind, um thermische und kalte Neutronen ohne wesentliche Verluste in entfernte Experimentiersäle zu leiten. Durch die schwache Biegung der Neutronenleiter gelingt es, thermische und kalte Neutronen verlustlos im Rohr weiterzuleiten, während die epithermischen und schnellen Neutronen und natürlich auch die γ-Strahlung geradlinig aus dem Bündel herausführen.

3.4.4.4. Neutronennachweis und -spektroskopie

3.4.4.4.1. *Nachweis.* Ein notwendiges Element jeder neutronenphysikalischen Anordnung ist der Detektor. Er hat die Aufgabe, Neutronen nachzuweisen und gegebenenfalls ihre Energie zu bestimmen. Die zum Nachweis für Neutronen benutzten Elementarprozesse sind die Auslösung von Kernreaktionen durch Neutronen und der Nachweis der Reaktionsprodukte (z. B. die ^{10}B (n,α)^{7}Li-Reaktion), die Messung der in speziellen Detektorfolien induzierten künstlichen Radioaktivität (Sondenmethode) und der Nachweis von elastisch gestreuten Rückstoßkernen (in erster Linie Protonen) in Ionisationskammern oder Szintillatoren. Das letztgenannte Verfahren eignet sich nur für Neutronen mit Energien > 10 keV.
Da die Energien der zu registrierenden Neutronen einen Bereich von 10 Größenordnungen überdecken, gibt es keine universellen Neutronendetektoren.

3.4.4.4.2. *Spektroskopie.* Die Bestimmung der Energieverteilung von Neutronen erfordert grundsätzlich andere experimentelle Methoden als die Spektroskopie geladener Teilchen. Möglich ist z. B. die *Spektroskopie der Rückstoßprotonen.* Bei gegebenem Beobachtungswinkel sind die Energie des Rückstoßkerns (z. B. Proton) und die Energie des zu analysierenden Neutrons korreliert. Außerdem ist die *Ausnutzung von Kernreaktionen* möglich. Der Nachweis der Neutronen geschieht dann in sogenannten Schwellendetektoren. Hierzu werden endotherme Reaktionen verwendet, welche erst oberhalb einer bestimmten Neutronenenergie ansprechen. Durch Kombination verschiedener Substanzen lassen sich Aussagen über die Energieverteilung der Neutronen am Ort dieser Sonden machen. Eine dritte Möglichkeit stellt die *Ausnutzung der Welleneigenschaften der Neutronen* dar. Wie bereits beschrieben (↑ 3.4.4.3.), kann die Braggreflexion zur Aussonderung und damit in Verbindung mit einem Detektor zur Spektroskopie von Neutronen bis zu etwa 10 eV verwendet werden. Weiterhin werden *Laufzeitmethoden* benutzt. Die Laufzeit entlang einer Strecke ist durch die Energie des Neutrons bestimmt. Für die Energiebestimmung benötigt man möglichst genau den Startzeitpunkt und den Zeitpunkt der Ankunft des Neutrons am Ende der Strecke. Die letztere Angabe liefert der Detektor, ein genauer Startzeitpunkt ist im allgemeinen durch die Pulsung der Neutronenquelle gegeben. Die gepulsten Neutronenquellen an Beschleunigern und die Impulsreaktoren sind deshalb mit Laufzeitstrecken ausgerüstet. Die am Impulsreaktor im VIK Dubna maximal verfügbare Laufzeitstrecke beträgt 1000 m. Steht ein Reaktor als Neutronenquelle zur Verfügung, so können Neutronenimpulse auch durch mechanische Unterbrechung des aus einem Kanal des Reaktors austretenden Neutronenstrahls erzeugt werden. Da bei gegebener Impulsdauer das Auflösungsvermögen der Anordnung $\Delta E/E$ mit der Länge der Laufstrecke verbessert wird, andererseits jedoch die Intensitätsparameter ungünstiger werden, verläuft die Entwicklung nach kurzen, aber intensiven Neutronenimpulsen sowie relativ kurzen Laufzeitstrecken. Hierbei bieten die Beschleuniger günstige Voraussetzungen, um Impulslängen von wenigen 10^{-9} s Dauer zu erzeugen.
Eines der Hauptanwendungsgebiete der Neutronenspektroskopie ist die Messung von Wirkungsquerschnitten als Funktion der Neutronenenergie. Diese Angaben sind insbesondere für die Projektierung von Kernreaktoren von großer Bedeutung.

3.4.4.5. Anwendungen der Neutronenbeugung und Neutronenstreuung. Die Neutronenbeugung gestattet Untersuchungen der Gitterstrukturen von Kristallen analog zur Röntgenstrukturanalyse. Obwohl sie gegenüber letzterer einen wesentlich höheren Aufwand erfordert, stellt sie eine unersetzliche Methode der Festkörperforschung dar. Ihre Bedeutung geht aus den Unterschieden der Eigenschaften der Neutronen und deren Wechselwirkung gegenüber der Röntgenstrahlung hervor.

Die Neutronen treten über die Kernkräfte mit den Atomkernen in Wechselwirkung. Die Streuung und auch die Absorption ist ein rein kernphysikalischer Vorgang. Ihr kohärenter Anteil vermittelt über die Bragg'sche Reflexion ein Bild von der Lage der Kerne und nicht, wie bei der Röntgenbeugung, von der Verteilung der Elektronen. Bei der Neutronenbeugung können also besonders leichte Kerne im Kristallgitter lokalisiert werden, welche für die Röntgenstrukturanalyse praktisch nicht zugänglich sind. Allerdings wirkt sich gerade bei Wasserstoff der geringe Anteil des kohärenten Streuquerschnitts nachteilig aus, so daß man oft gezwungen ist, im Gitter den Wasserstoff (wenigstens an den interessierenden Gitterplätzen) durch Deuterium zu ersetzen.

Eine einzigartige Anwendung der Neutronenbeugung beruht auf der Wechselwirkung zwischen dem magnetischen Moment des Neutrons und dem magnetischen Moment der Elektronenhülle der Atome. Damit ist die magnetische Struktur der immer mehr an Bedeutung gewinnenden magnetischen Werkstoffe sowie der für die Grundlagenforschung interessanten Antiferromagnetika bestimmbar.

Ein weiteres Charakteristikum der Neutronenbeugung ist die im Vergleich zur Röntgenstrahlung geringe Absorption der Neutronen. Die damit verbundene Erfassung tiefer liegender Probenbereiche erlaubt die zerstörungsfreie Untersuchung der Texturen von Werkstoffen.

Die inelastische Neutronenstreuung vermittelt Kenntnisse über die Bewegungen der Atome und Moleküle in Festkörpern und Flüssigkeiten. Thermische Neutronen sind besonders gut geeignet zur Unterscheidung der Schwingungen der Atome um ihre Gleichgewichtslagen im Gitter, da sowohl Impuls als auch Energie der Neutronen die gleiche Größenordnung besitzen wie Impuls und Energie der Gitterbausteine. Daher lassen sich die mit der Anregung dieser Schwingungen verbundenen Änderungen von Impuls und Energie der Neutronen gut nachweisen. Nach den Gesetzen von Impuls- und Energieerhaltung wird dabei zwischen den Neutronen und den Atomkernen des Gitters in elastischen Stößen kinetische Energie ausgetauscht. Gegenüber dem Gitter als Ganzes entspricht das einer inelastischen Streuung der Neutronen bei gleichzeitiger Veränderung des Anregungszustands der Gitterschwingungen. Analoge Untersuchungen sind mit Röntgenstrahlen nicht möglich, da wegen der sehr kleinen Impulse der Röntgenquanten praktisch kein Impulsaustausch mit dem schwingenden Gitter möglich ist.

Die Methode der inelastischen Streuung ist nicht nur auf die Untersuchung der Gitterschwingungen beschränkt, sie läßt sich auf ungeordnete Substanzen wie amorphe Festkörper, Flüssigkeiten, Polymere oder biologische Substanzen zur Untersuchung dynamischer Prozesse anwenden. Die geringe kinetische Energie der thermischen Neutronen garantiert dabei eine minimale Einwirkung auf das Probenmaterial, so daß selbst biologische Moleküle im Lebensprozeß oder katalytische Reaktionen verfolgt werden können.

3.4.4.6. Abschirmung und Filterung von Neutronen. Hierzu werden Substanzen verwendet, welche einen großen Absorptionsquerschnitt für Neutronen haben. Da die Querschnitte stark von der Energie abhängen, gibt es keine universelle Abschirmung.

Für thermische Neutronen eignen sich Cd, B oder Li bereits in ihren natürlichen Isotopengemischen, obwohl immer nur eines ihrer Isotope hohe Wirkungsquerschnitte zeigt. Aufgrund einer starken Resonanz bei 0,176 eV absorbiert Cd praktisch alle Neutronen unterhalb 0,4 eV, auch wenn es in einer Stärke von nur einigen zehntel Millimetern vorliegt. Für Resonanz- und schnelle Neutronen ist dieses Material aber noch durchlässig. Dadurch entsteht eine Art Filterwirkung, mit der thermische und epithermische Neutronen voneinander getrennt werden können.

Schnelle Neutronen können nicht so einfach abgeschirmt werden, da die Wirkungsquerschnitte aller Materialien für sie nur in der Nähe von $1\ b = 10^{-28}\ m^2$ liegen. Eine relativ wirksame Abschirmung wird durch die Kombination von wasserstoffreichem Material und Absorbermaterial erreicht, wie etwa Paraffin und Borsäure. Der Wasserstoff sorgt dabei für die Abbremsung der Neutronen in den thermischen Bereich, in welchem das Absorbermaterial wirksam wird. Für die Thermalisierung werden Paraffin- oder Wasserschichten von 10 bis 20 cm Stärke benötigt.

Eine andere Möglichkeit besteht in der Ausnutzung der inelastischen Streuung, welche besonders im MeV-Bereich zu einer wirksameren Abbremsung führt als die elastische Streuung an Wasserstoff. Der in Kernanlagen verwendete Schwerbeton besteht deshalb zu einem wesentlichen Anteil aus Eisenschrott. Ihm sind gleichzeitig borhaltige Materialien beigesetzt.

Während die Filterung von schnellen Neutronen aus einem breiten Neutronenspektrum bereits durch Cd-Folien realisiert werden kann, ist das

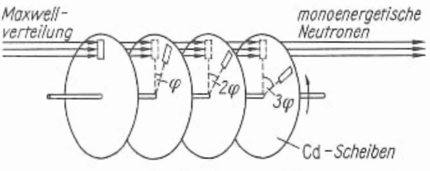


Abb. 3.4.4-4 Prinzip eines mechanischen Neutronenfilters

Herausfiltern thermischer Neutronen problematisch. Benutzt werden mechanische Selektoren aus schnell rotierenden Scheiben oder kompakten Wellen mit eingefrästen Schlitzen, welche nur Neutronen eines bestimmten Geschwindigkeitsbereichs den Durchgang gestatten (Abb. 3.4.4-4).

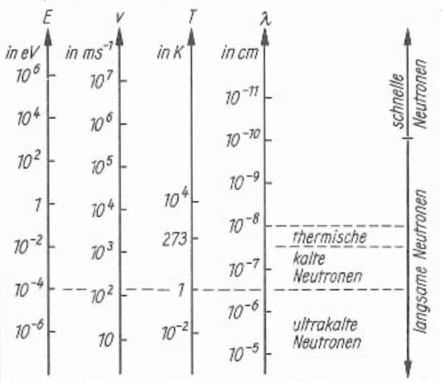


Abb. 3.4.4-5 Beziehungen zwischen Energie, Geschwindigkeit, Wellenlänge und Temperatur für thermische, kalte und ultrakalte Neutronen

Für Arbeiten mit thermischen Neutronen stehen an Forschungsreaktoren die sogenannten thermischen Säulen zur Verfügung. Sie bestehen aus großen Graphitblöcken, welche sich in entsprechenden Öffnungen der Betonabschirmung des Reaktors befinden. Dadurch wird das Verhältnis von thermischen zu schnellen Neutronen stark angehoben. Nachteilig ist dabei jedoch auch ein wesentlicher Verlust an thermischen Neutronen. Deshalb sind unter Umständen selektivere Filtermethoden erforderlich.

Ein anderes Prinzip der Filterung basiert auf der Streuung von Neutronen an polykristallinen Substanzen. Sie gestatten nur Neutronen oberhalb einer Grenzwellenlänge, welche durch den Gitterabstand der Kristallstruktur bestimmt ist, den Durchgang, während schnellere Neutronen unter den vielen Orientierungen im Polykristall sicher eine Fläche finden, an der sie aus dem Strahl herausgestreut werden. Die Grenzwellenlängen liegen zwischen $4 \cdot 10^{-10}$ m für Be und $7 \cdot 10^{-10}$ m für Graphit und Bi, d. h. die Filter sind für Neutronen unterhalb $\approx 10^{-3}$ eV (kalte Neutronen) durchlässig (Abb. 3.4.4-5).

Die eleganteste Methode stellt jedoch die Herausführung von thermischen und kalten Neutronen durch Neutronenleiter (↑ 3.4.4.3.3.) aus dem allgemeinen Feld epithermischer und schneller Neutronen sowie der γ-Strahlung dar.

3.4.4.7. Ultrakalte Neutronen. Entsprechend ihrer mittleren Geschwindigkeit bei Zimmertemperatur werden Neutronen von etwa 0,025 eV als thermisch bezeichnet (↑ Abb. 3.4.4-5). Dementsprechend haben kalte Neutronen Energien unterhalb 10^{-2} eV und ultrakalte Neutronen unterhalb 10^{-4} eV. Letzteren entspricht, falls sie im Maximum einer Maxwellverteilung liegen sollten, eine Temperatur von weniger als 1 K (Temperatur des Neutronengases im thermischen Gleichgewicht). Von besonderem Interesse sind ultrakalte Neutronen mit Energien von etwa 10^{-7} eV, die 1969 am Forschungsreaktor des Kurtschatow-Instituts Moskau nachgewiesen wurden. Ihre Geschwindigkeit beträgt nur noch wenige ms^{-1}, die entsprechende de-Broglie-Wellenlänge 10^{-7} m.

Derartig niederenergetische Neutronen kommen in der Geschwindigkeitsverteilung eines Moderators bei Zimmertemperatur nur in geringer Anzahl vor. Ihr Verhältnis zur Gesamtneutronenzahl beträgt etwa 10^{-9}. Man ist deshalb gezwungen, mit wenigen Neutronen zu experimentieren. Dabei kommt die Tatsache zu Hilfe, daß schnellere Neutronen an den Innenwänden der Neutronenleiter und Gefäße nicht reflektiert werden und somit nach außen entweichen. Ultrakalte Neutronen erleiden dagegen an verschiedenen Materialien bei beliebigem Auftreffwinkel Totalreflexion. Sie sind also in einem Gefäß, welches innen mit dem entsprechenden Material ausgekleidet ist, einschließbar (Abb. 3.4.4-6a) und können nur durch ihren eigenen radioaktiven Zerfall oder durch unelastische Prozesse mit der Wandung verlorengehen. Bei der experimentellen Überprüfung der theoretisch ermittelten Speicherzeit ergaben sich jedoch bisher Diskrepanzen, welche noch nicht geklärt werden konnten.

Entsprechend ihrer geringen Geschwindigkeit können ultrakalte Neutronen im Gravitationsfeld der Erde nur geringe Höhendifferenzen

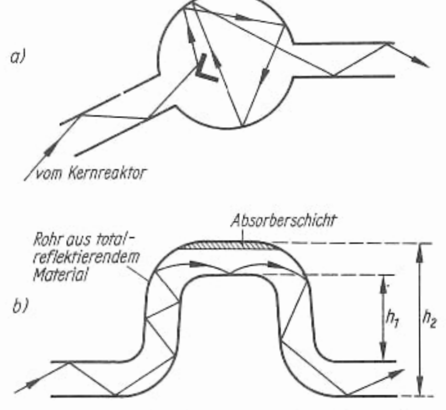


Abb. 3.4.4-6 a) Einschluß ultrakalter Neutronen am Kernreaktor, b) Ausblendung monochromatischer ultrakalter Neutronen

überwinden. Das ergibt eine bequeme Methode der Ausblendung bestimmter Energiebereiche. Schickt man die Neutronen durch einen Leiter der in Abb. 3.4.4-6b angegebenen Form, so begrenzen die Höhen h_1 und h_2 den Durchlaßbereich. Es werden Energieschärfen von einigen 10^{-9} eV erreicht.

Zur Erhöhung der Zahl der verfügbaren ultrakalten Neutronen kann man schnellere Neutronen in einem vertikalen Neutronenleiter durch das Schwerefeld der Erde abbremsen. Als weitere Methoden werden neuerdings tiefgekühlte Moderatoren sowie Neutronenturbinen verwendet. Bei der ersten Methode wird das Maxwell'sche Geschwindigkeitsspektrum generell zu kleineren Energien hin verschoben, so daß die im interessierenden Bereich vorhandene Neutronenzahl relativ zur Gesamtzahl wächst. Bei der zweiten Methode wird kinetische Energie der Neutronen teilweise bei der Totalreflexion an die bewegten Schaufeln einer Turbine abgegeben.

3.5. Kernenergetik

3.5.1. Reaktorphysik

3.5.1.1. Kernspaltung. Ein Maß für die Stabilität eines Atomkerns ist seine Bindungsenergie, die bezogen auf ein einziges Nukleon die Energie darstellt, die man aufbringen muß, um ein Nukleon vom Kern abzutrennen. Sie ist für die leichtesten und schwersten Kerne kleiner als für die mittleren (↑ 3.3.2.2.). Der physikalische Grund für die geringere Stabilität der schwersten Kerne liegt in der großen Zahl von Protonen im Kern, deren elektrostatische Abstoßung den anziehenden Kernkräften merkbar entgegenwirkt, was schließlich zum Abbruch des Periodensystems der Elemente führt.

Obgleich die Spaltung eines schweren Kerns in zwei mittlere Kerne mit einer Energiefreisetzung verbunden wäre, ist eine *spontane Kernspaltung* außer bei Transuranen eine sehr unwahrscheinliche und daher sehr seltene Reaktion. Die Spaltung in zwei leichtere Bruchstücke erfordert nämlich die Überwindung einer Energieschwelle (↑ 3.4.1.3.1.), bevor der Spaltungsprozeß ablaufen kann. Die zur Überwindung dieser Schwelle notwendige Energie E_w heißt *Aktivierungsenergie*. Sie nimmt mit wachsender Massenzahl A des Kerns ab. Für $A > 250$ ist $E_w = 0$, für $A \approx 230$ ist $E_w < 10$ MeV und wächst für $A \approx 130$ auf etwa 100 MeV an. Demnach lassen sich nur die schwersten Kerne relativ leicht spalten. Die dabei freiwerdende Energie ist die eigentliche wirtschaftlich zu nutzende Kernenergie.

Im interessierenden Energiebereich kommen als Träger der Aktivierungsenergie praktisch nur neutrale Teilchen, also γ-Quanten und Neutronen in Frage. Diese können jedoch durch elektrische Felder nicht beschleunigt werden, so daß den Kernen auf diese Weise nur eine Aktivierungsenergie von maximal 10 MeV zugeführt werden kann. Damit bleibt die praktisch zu nutzende Spaltreaktion auf Kerne mit $A \gtrsim 230$ beschränkt. Die Erzeugungsrate geeigneter γ-Quanten ist jedoch praktisch so gering, daß die bewirkten Spaltreaktionen selbst für Anlagen kleinster Leistungen nicht genutzt werden können. So bleiben derzeit die Neutronen die einzigen Träger der für eine Spaltung schwerer Kerne notwendigen Aktivierungsenergie.

Dringt ein Neutron in einen schweren Kern ein, so entsteht ein Zwischenkern, der ein schwereres Isotop des Ausgangskerns darstellt. Dieser Zwischenkern ist jedoch um die kinetische Energie E_k des Neutrons und seine freiwerdende Bindungsenergie $E_B^{(n)}$ angeregt. Wird diese Anregungsenergie als γ-Quant wieder abgestrahlt, so liegt ein *Strahlungseinfang* vor. Ist aber $E_k + E_B^{(n)} > E_w$, so kann die Spaltung des Kerns eintreten.

Bei U-235 und Pu-239 reicht bereits die Bindungsenergie des Neutrons zur Kernspaltung aus, so daß bereits Neutronen mit kinetischen Energien $E_k \approx 0{,}025$ eV, die sogenannten *thermischen Neutronen*, Spaltungen dieser Kerne auslösen können. Dagegen sind für die Spaltung von U-238-Kernen schnelle Neutronen mit $E_k \geqq 1{,}5$ MeV erforderlich (Tab. 3.5.1-1). Das liegt daran, daß Kerne mit gerader Massenzahl stabiler sind als die mit ungerader (↑ 3.3.2.2.4.). U-233, U-235 und Pu-239 sind die leichtesten der schweren Nuklide, die durch thermische Neutronen gespalten werden können.

Tab. 3.5.1-1 Aktivierungsenergie E_W einiger schwerer Kerne in MeV; E_k kinetische Energie des Neutrons, E_B Bindungsenergie

	U-235	U-238	Pu-239
E_W	6,5	7,0	5,1
E_B	6,8	5,5	5,1
E_k	0	1,5	0

3.5.1.2. Kettenreaktion. Um die Spaltung schwerer Kerne durch Neutronen wirtschaftlich nutzen zu können, braucht man kontinuierlich eine große Anzahl von Kernspaltungen. Eine Radium-Beryllium-Neutronenquelle mit 10^{10} Neutronen pro Sekunde würde, abgesehen von ihren hohen Kosten, im günstigsten Falle eine Spaltrate (Anzahl der Spaltakte pro Sekunde) liefern, die einer Leistung von nur einigen Zehntel Watt entspräche. Es werden aber

bei einer Spaltung schwerer Kerne im Mittel zwei bis drei Neutronen pro Kern herausgeschleudert, sogenannte *Spaltneutronen*, die ihrerseits weitere Kernspaltungen hervorrufen können. Wenn ein Reaktor so konstruiert wird, daß die durch Abwanderung aus ihm bzw. durch Strahlungseinfang entstehenden Verluste von Spaltneutronen so eingeschränkt werden, daß im Mittel jede Kernspaltung über die Spaltneutronen mindestens eine weitere auslöst, kann die so unter Kontrolle verlaufende Kettenreaktion für Reaktoren im Prinzip beliebig hoher Leistungen genutzt werden (s. a. 3.5.1.4.).

3.5.1.3. Neutronenreaktionen im Reaktor. Die bei einer Kernspaltung freiwerdenden Spaltneutronen haben eine kinetische Energie von 1,5 bis 10 MeV. Sie können an den im Reaktor vorhandenen verschiedenen Kernen z. B. von Brennmaterial, Moderator, Konstruktionsmaterial oder Kühlsubstanz, gestreut oder durch sie absorbiert werden. Die Spaltreaktion durch ein Neutron ist natürlich die wichtigste Reaktion im Reaktor.

Ob und wie ein Neutron mit einem Kern reagiert, wenn es in seine Nähe gelangt, wird durch den Wirkungsquerschnitt σ (↑ 3.4.2.2.) charakterisiert, der die Wahrscheinlichkeit der entsprechenden Reaktion bestimmt und experimentell ermittelt wird. Der Wirkungsquerschnitt σ_i für eine bestimmte Reaktion i ist von der Kernart, aber auch von der kinetischen Energie der Neutronen abhängig. Da die Verweilzeit in Kernnähe bei langsamen Neutronen größer ist als bei schnellen, ist auch die Spaltreaktion für langsame Neutronen wahrscheinlicher als für schnelle, wenn der Kern bereits durch langsame Neutronen spaltbar ist. Es gilt $\sigma_f \sim v^{-1}$, wobei σ_f der Spaltquerschnitt und v die Neutronengeschwindigkeit ist. Demnach ist es effektiver, die schnellen Spaltneutronen (1 bis 10 MeV) durch Streuprozesse auf niedrigste, also auf *thermische Energien* von im Mittel 0,025 eV abzubremsen, bevor sie mit einem spaltbaren Kern zusammenstoßen. Dazu werden dem spaltbaren Brennmaterial sogenannte *Moderatoren* beigemengt, die die gewünschten Bremsvorgänge bewirken. Um dabei Neutronenverluste durch Einfang im Moderator, in den Konstruktionsmaterialien oder auch im Brennstoff selbst zu vermeiden bzw. minimal zu halten, müssen Moderator und Konstruktionsmaterialien kleine Absorptionsquerschnitte für Neutronen haben und der Bremsvorgang so rasch wie möglich, d. h. durch nur wenige Stöße mit den Moderatorkernen, erfolgen. Während des Bremsvorgangs durchlaufen die Neutronen Energien von einigen MeV bis 0,025 eV. Das Gebiet zwischen 1000 eV und 1 eV ist dabei besonders gefährlich, weil dort an den meisten Kernen bei ganz bestimmten diskreten Neutronenenergien mit hoher Wahrscheinlichkeit ein sogenannter *Resonanzeinfang* erfolgt. Dementsprechend hängt der Einfangquerschnitt σ_{abs} sehr stark von der Neutronenenergie ab. Ist die Neutronenenergie etwas größer oder kleiner als die *Resonanzenergie*, so zeigt der Kern nur geringe Bereitschaft zur Neutronenabsorption. Es ist daher wichtig, daß die Neutronen beim Bremsvorgang dieses Resonanzgebiet der Energie möglichst überspringen, also bei einem Stoß mit einem Moderatorkern möglichst viel Energie verlieren.

Die *elastische Streuung* ist die einfachste Neutronenreaktion. Das Neutron prallt mit einem Kern zusammen und tauscht mit ihm Energie und Impuls aus. Dabei wird der Kern nicht angeregt. Wichtig für die Reaktortechnik ist, daß beim Stoßprozeß ein Teil der Neutronenenergie auf den Kern als Ganzes übertragen wird, und zwar um so mehr, je leichter der Kern ist. Demnach muß ein zur Neutronenbremsung eingesetzter Moderator aus leichten Kernen bestehen, die zudem einen kleinen Einfangquerschnitt für Neutronen besitzen.

3.5.1.4. Spalt-, Brut- und Kernbrennstoffe. Eigentliche Spaltstoffe sind nur die Nuklide U-233, U-235 und Pu-239. Sie können bereits durch thermische Neutronen gespalten werden. Jedoch kommt nur das U-235 in der Natur vor; es ist zu 0,71 % im natürlichen Uranium enthalten, welches im wesentlichen aus U-235 und U-238 im Mischungsverhältnis 1 : 140 besteht. In *angereichertem Uranium* ist der Anteil der U-238-Kerne durch technisch aufwendige Isotopentrennverfahren (↑ 3.3.1.4.) künstlich vermindert. U-238 ist nur sehr schlecht spaltbar, hat aber die Fähigkeit, sowohl schnelle als auch thermische Neutronen einzufangen. Dabei entsteht das Unraniumisotop U-239, das durch *radioaktive Umwandlung* in das relativ stabile, aber in der Natur nicht vorkommende Pu-239 übergeht:

$$^{238}_{92}\mathrm{U}(\mathrm{n},\gamma)^{239}_{92}\mathrm{U} \xrightarrow[\beta^-]{2{,}3\,\mathrm{d}} {}^{239}_{93}\mathrm{Np} \xrightarrow[\beta^-]{23\,\mathrm{min}} {}^{239}_{94}\mathrm{Pu}.$$

Ähnlich ist es mit dem in der Natur genügend vorkommenden Th-232. Es ist selbst kaum spaltbar; seine Kerne verwandeln sich jedoch durch Neutroneneinfang in das gut spaltbare Uraniumisotop U-233:

$$^{232}_{90}\mathrm{Th}(\mathrm{n},\gamma)^{233}_{90}\mathrm{Th} \xrightarrow[\beta^-]{23{,}5\,\mathrm{min}} {}^{233}_{91}\mathrm{Pa} \xrightarrow[\beta^-]{27{,}4\mathrm{d}} {}^{233}_{92}\mathrm{U}.$$

Bringt man somit neben dem Spaltmaterial auch die Isotope U-238 und Th-232 im Reaktor unter, so können auf diese Weise durch überschüssige Neutronen neue Spaltstoffe erzeugt werden. Dieser Vorgang heißt *Brüten*; die entsprechenden Reaktoren nennt man *Brüter*. Charakteristisch für einen Brüter ist seine *Brutrate*, d. h. die Anzahl der erzeugten Spaltstoffkerne pro verbrauchten Spaltstoffkern. Da die

Brutstoffe U-238 und Th-232 – im Unterschied zu den sehr begrenzten Vorkommen des Spaltstoffs U-235 – in der oberen Erdkruste in ausreichenden Mengen vorkommen, wird die Kernenergie erst durch Brüter mit einer Brutrate > 1 zur wahrscheinlichen Energiequelle der Zukunft. Reaktoren arbeiten nicht mit reinem Spaltstoff, sondern mit *Kernbrennstoff*, einer Substanzmischung, in der Spaltstoff homogen verteilt ist, z. B. Uranium jeglichen Anreicherungsgrads, Uranium-Thorium-Gemische, die als Metalle, Legierungen, Oxide, Carbide vorliegen. Es können auch Reaktoren mit Uraniumsalzlösungen betrieben werden. Natururanium ist der billigste Brennstoff. Sein Neutronenüberschuß zum Ausgleich unvermeidlicher Verluste ist jedoch gering. Die meisten Reaktoren werden daher mit angereichertem Uranium betrieben.

3.5.1.5. Kritische Größen. Um eine *stationäre Spaltkette* zu verwirklichen, muß jede Neutronengeneration die gleiche Anzahl von Spaltungen hervorrufen wie die vorhergehende. Es muß eine genügend große Menge an Spaltstoff in der Anordnung vorhanden sein, damit die Neutronenverluste durch Entweichen in die Umgebung und durch Neutronenabsorption in der Anordnung durch überschüssige Spaltneutronen aufgewogen werden. Die Mindestmenge an Spaltstoff, die in einer vorgegebenen Anordnung eine stationäre Kettenreaktion aufrechterhält, heißt *kritische Menge*. Das dazu notwendige Mindestvolumen eines Reaktors heißt *kritisches Volumen*. Überschreitungen der kritischen Größen führen zum exponentiellen Anwachsen der Kettenreaktion und damit der erzeugten Leistung. Solche Anordnungen heißen überkritisch. Bleibt man dagegen unter den kritischen Maßen, so kommt eine eingeleitete Kettenreaktion mit Sicherheit exponentiell zum Erliegen. Die Aufgabe der Reaktorphysik ist es, Form, Größe und Materialverteilung für einen Reaktor so zu bestimmen, daß er im kritischen Regime, also stationär, arbeitet und seine Leistung geregelt werden kann.

3.5.1.6. Moderatoren. Die heute zur praktischen Energieerzeugung eingesetzten Reaktoren arbeiten mit thermischen Neutronen. Das wird sich in den nächsten Jahrzehnten nicht ändern, und selbst dann, wenn zu Beginn des nächsten Jahrhunderts zunehmend der schnelle Brüter zum Einsatz kommt, werden die thermischen Reaktoren noch auf lange Sicht eine große Rolle spielen. Im thermischen Reaktor muß dem Brennstoff eine Moderatorsubstanz beigefügt werden, an deren Kernen die Spaltneutronen durch Streuung auf thermische Energien abgebremst werden (↑ 3.5.1.3.).

Der Energieverlust pro Stoß wird in der Reaktorphysik nicht durch die Differenz $\Delta E = E' - E''$ der Neutronenenergien vor und nach dem Stoß charakterisiert, sondern durch die Differenz der Logarithmen dieser Größen: $\ln E' - \ln E'' = \ln E'/E''$. Der über viele Stöße gemittelte logarithmische Energieverlust pro Stoß $\xi = \overline{\ln E'/E''}$ erweist sich als nur von der Massenzahl A der Kerne abhängig und kann für $A > 2$ genähert durch $\xi \approx 2/(A + 2/3)$ ausgedrückt werden. Er ist für Wasserstoffkerne ($A = 1$) am größten. Als *Bremsvermögen* eines Moderators wird die Größe $n\sigma_{el}\xi \equiv \Sigma_{el}\xi$ bezeichnet, wo n die Anzahl der Moderatorkerne pro Volumeneinheit bedeutet. $\Sigma_{el} = n\sigma_{el}$ wird in der Reaktorphysik makroskopischer Streuquerschnitt genannt. Sein reziproker Wert $l_{el} = \Sigma_{el}^{-1}$ hat die Bedeutung des mittleren Abstands zwischen den Stößen, bei denen die Neutronen gestreut werden (mittlere freie Streuweglänge). Gewöhnliches Wasser hat das größte Bremsvermögen. Dennoch ist es nicht der effektivste Moderator, da es Neutronen stark absorbiert, also einen großen Absorptionsquerschnitt σ_{abs} für Neutronen besitzt. Die tatsächliche Effektivität eines Moderators wird durch das *Bremsverhältnis* $\xi\Sigma_{abs}/\Sigma_{abs}$ mit $\Sigma_{abs} = n\sigma_{abs}$ bestimmt. Danach ist schweres Wasser D_2O, in dem die Wasserstoffkerne aus je einem Proton und einem Neutron bestehen, der effektivste Moderator. Dann folgen reiner Graphit, Beryllium und leichtes Wasser (Tab. 3.5.1-2). Letzteres ist als Moderator am billigsten, schweres Wasser dagegen am teuersten. Mit schwerem Wasser moderierte Reaktoren haben die kleinsten kritischen Dimensionen.

Tab. 3.5.1-2 Bremseigenschaften verschiedener Moderatoren M; ξ mittlerer logarithmischer Energieverlust; $\xi\Sigma_{el}$ Bremsvermögen, $\xi\Sigma_{el}/\Sigma_{abs}$ Bremsverhältnis

M	ξ	$\xi\Sigma_{el}$ in cm^{-1}	$\xi\Sigma_{el}/\Sigma_{abs}$
H_2O	0,927	1,53	70
D_2O	0,508	0,170	5200
Be	0,206	0,178	162
C	0,158	0,061	235

In homogenen Reaktoren ist der Moderator mit dem Brennstoff gleichmäßig vermischt. Um den Anteil der in U-238 resonanzabsorbierten Neutronen herabzusetzen, wird in *heterogenen Reaktoren* (Abb. 3.5.1-3) der Brennstoff in Stabform in den Moderator eingeführt, wobei der Stabdurchmesser kleiner ist als der Abstand zwischen den Stäben. Auf diese Weise verlassen die schnellen Spaltneutronen nach ein bis zwei Stößen den Brennstoffstab und werden im reinen Moderator über das gefährliche Resonanzgebiet (1000 bis 1 eV) auf thermische Energien abgebremst, bevor sie den nächsten Brennstoffstab erreichen. Wie sehr die durch Reso-

nanzabsorption bewirkten Neutronenverluste dadurch herabgesetzt werden, ersieht man daraus, daß ein mit Natururanium betriebener graphitmoderierter homogener Reaktor überhaupt nicht kritisch wird, während der heterogene Typ zur Energieerzeugung eingesetzt wird.

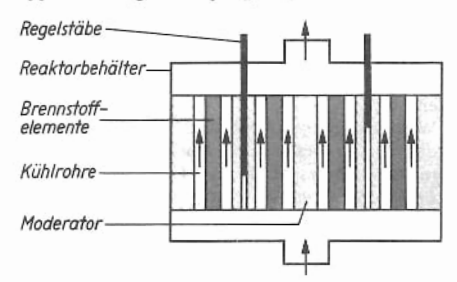


Abb. 3.5.1-3 Schema eines heterogenen Reaktors

In diesem Zusammenhang ist die Frage von Interesse, ob nicht auch zufällige Anhäufungen von Uraniummineralien in der Natur zur Auslösung von Kettenreaktionen geführt haben können. Nachdem dieses Problem mehrfach theoretisch ohne endgültige Klärung diskutiert worden war, hat man vor einigen Jahren tatsächlich in einem Minengebiet von Gabun eine Formation aufgefunden, die mit großer Wahrscheinlichkeit als ein ausgebrannter natürlicher Kernreaktor anzusehen ist (Tafel 6).

3.5.1.7. Spaltprodukte. Bei der Kernspaltung entstehen als Spaltprodukte zwei in der Regel ungleiche mittelschwere Kerne, beispielsweise $^{233}_{92}U \rightarrow ^{94}_{38}Sr + ^{139}_{54}Xe$.
Es treten insgesamt über 200 Nuklide der Elemente von Zink (Z = 30) bis Therbium (Z = 65) mit den Massenzahlen $72 \leqq A \leqq 165$ auf (Abb. 3.5.1-4). Ihre kinetische Energie stellt den Hauptanteil der freiwerdenden Kernenergie dar. Sie sind neutronenreich und daher in der Regel instabil. Durch β-Zerfall verwandeln sie sich mit den verschiedensten Halbwertszeiten in stabile Elemente, beispielsweise

$$^{137}_{53}I \xrightarrow[\beta^-]{23\,s} {}^{137}_{54}Xe \xrightarrow[\beta^-]{3{,}4\,min} {}^{137}_{55}Cs \xrightarrow[\beta^-]{33\,a} {}^{137}_{56}Ba.$$

Sowohl der Spaltakt, als auch die β-Zerfälle der Spaltprodukte sind von γ-Strahlung begleitet, was nicht nur hohe biologische Sicherheitsanforderungen an den Reaktorbetrieb stellt, sondern auch hohe Forderungen an die radioaktive Belastbarkeit der Konstruktionswerkstoffe.

Die Gesamtheit der Spaltprodukte stellt ein heterogenes, hochradioaktives Gemisch verschiedenartigster Elemente dar, das sich schwierig transportieren, aufbereiten und lagern läßt. Wegen der langen Zerfallszeit von einigen Jahrzehnten bis Jahrhunderten mancher Spaltprodukte stellt ihre sichere Aufbewahrung, die ein Eindringen auch von Spuren dieser Substanzen in den Biozyklus verhindern muß, eines der schwierigsten Probleme der in großen Maßstäben angewandten Kernenergetik dar.

Viele Spaltprodukte haben einen großen Absorptionsquerschnitt σ_{abs} für thermische Neutronen und stören daher empfindlich den Neutronenhaushalt. Dieser Effekt wird *Vergiftung* bzw. *Verschlackung* genannt. Ein genau kritisch ausgelegter Reaktor käme nach kurzer Zeit zum Stillstand, da außerdem der Spaltstoffanteil mit zunehmender Betriebsdauer abnimmt. Ein für längere Betriebszeit vorgesehener Reaktor muß daher überkritisch ausgelegt werden. Dieser Überschuß an „Reaktivität" wird durch *Regelstäbe* kompensiert, die aus einem starken *Neutronenabsorber*, z. B. Cadmium, bestehen und in den Reaktor so tief eingesenkt werden, bis er kritisch arbeitet. Bei zunehmender Verschlackung und Spaltstoffverbrauch werden die Stäbe langsam ausgefahren bis schließlich die ganze „Überschußreaktivität" verbraucht ist. Die Brennelemente müssen dann ausgewechselt werden, obgleich sie noch etwa 50% unverbrauchten Spaltstoff enthalten. Die chemische Aufbereitung des Brennstoffs stellt ein kompliziertes und kostspieliges Verfahren dar, bei welchem Brennstoffverluste unvermeidlich sind. In derzeitigen thermischen Reaktoren wird somit der nukleare Energievorrat nur teilweise genutzt.

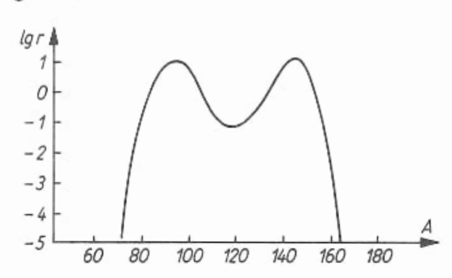


Abb. 3.5.1-4 Relativer Anteil r (in Prozent) der Spaltprodukte verschiedener Massenzahlen A an der Gesamtmenge von Folgekernen bei der thermischen U-235-Spaltung

Im Reaktor gewinnt man je 1,25 g verbrauchten Spaltstoff rund $24 \cdot 10^3$ kWh thermische Energie. Wird diese Wärmemenge in einem Kernkraftwerk mit einem Wirkungsgrad von 25% in elektrische Energie umgesetzt, so verbraucht ein Kraftwerk mit etwa 1000 MW elektrischer Leistung rund 4 kg Spaltstoff pro Tag. Ein konventionelles Kraftwerk gleicher Leistung braucht rund 40000 t Braunkohle je Tag.

3.5.2. Technik von Kraftwerksreaktoren

3.5.2.1. Reaktorkühlung. Die in den Brennelementen erzeugte Wärme gelangt durch Wärmeleitung an deren Oberfläche und wird dort an das Kühlmittel übertragen, welches die Brennstäbe in Längsrichtung umspült. Durch die ständige Wärmeaufnahme steigt die Temperatur eines von unten eintretenden Kühlmittels mit der Höhe im Reaktor. Die im Innern des Brennelements und entlang seiner Oberflächenhülle maximal zulässigen Temperaturen sind werkstoffbedingt. Die Wärmeleitfähigkeit des Kühlmittels, seine Wärmekapazität, seine Geschwindigkeit, die Geometrie der Kühlrohre müssen eine Überschreitung dieser Grenztemperaturen verhindern. Als Kühlmittel werden praktisch eingesetzt: Gase wie CO_2 und Helium, leichtes und schweres Wasser, Wasserdampf, organische Flüssigkeiten, flüssige Metalle wie Natrium. Im *Druckwasserreaktor* steht das Kühlmittel H_2O unter hohem Druck, um das Sieden zu vermeiden; im *Siedewasserreaktor* siedet es teilweise.
Das Kühlmittel transportiert die im Reaktor erzeugte Wärme in der Regel zu einem *Wärmetauscher*, in dem sie an ein Arbeitsmittel eines zweiten Kreislaufs abgegeben wird. Diese Trennung in zwei Kreisläufe beschränkt die Strahlenwirkung des radioaktiv verseuchten Kühlmittels und auch die bei Havarien eintretenden Strahlenschäden auf den Primärkreislauf, der einer besonderen Konstruktion und Überwachung unterliegt. Die weitere Überführung der Wärmeenergie in Elektroenergie verläuft wie in konventionellen Kraftwerken.
3.5.2.2. Werkstoffe, Konstruktionselemente, Anlagen. Die Brennstoffstäbe sind von einer metallischen Hülle, meist Aluminium, rostfreier Stahl, Zirconium, überzogen. Dieses sogenannte *Canning* hat die Aufgabe, die radioaktiven Spaltprodukte im Brennstoff zurückzuhalten, damit sie nicht in das Kühlwasser geraten. Die Anforderungen an diese metallischen Umhüllungen stellen die Techniker vor recht schwierige Probleme. Das Canningmaterial muß bei möglichst hohen Temperaturen mechanisch widerstandsfähig sein, eine gute Wärmeleitfähigkeit besitzen, gegenüber Neutroneneinwirkung, die gewöhnlich das Kristallgitter des Metalls zerstört, beständig sein und nur wenig Neutronen absorbieren. Es darf aber auch keine Verunreinigungen an fremden Kernen, die durch Bestrahlung zu langlebigen radioaktiven Isotopen führen, besitzen. Diese hohen Anforderungen an die *nukleare Reinheit* des Canningmaterials haben zur Entwicklung einer Zirconiummetallurgie geführt, da Zirconium für gasdichte Umhüllungen der Brennelemente besonders geeignet ist. Auch der Graphit kann wegen des gewünschten Reinheitsgrads nicht aus Naturgraphit gewonnen, sondern muß künstlich aus Kohlenwasserstoffen hergestellt werden. Breite Anwendung als Werkstoff für Bauelemente innerhalb des ersten Kreislaufs findet zunehmend nichtrostender Stahl.
Das wichtigste Element des Primärkreislaufs ist der *Reaktorbehälter*, der die Brennelemente, den Moderator, die Regelstäbe, die Rohrleitungen und die Umwälzpumpen für das Kühlmittel und, falls vorhanden, den Wärmetauscher enthält.
Weitere Anlagen sind: das *Druckhaltesystem*, das den vorgesehenen Druck des Kühlmittels in den vorgeschriebenen Grenzen hält; das *Lecksuchsystem* zum Nachweis defekter Brennelemente; das *Abgas-* und *Abwassersystem*, das die anfallenden radioaktiven Flüssigkeiten und Gase sammelt, transportiert und je nach dem Grad ihrer Radioaktivität speichert oder kontrolliert in die Umgebung abläßt; das *Kühlwasserreinigungssystem*, das ständig einen Teil des Kühlmittels von seinen radioaktiven Bestandteilen befreit; die *Beschickungseinrichtung* mit der, in Abhängigkeit vom Reaktortyp, kontinuierlich oder in bestimmten Zeitabständen die Entnahme, Beladung oder Umladung von Brennstoffkassetten durchgeführt wird; das *Lagerbecken*, in dem abgebrannte Brennstoffelemente gelagert werden, bis ein Teil der Radioaktivität abgeklungen ist. Daneben sind die *Strahlenschutzeinrichtungen*, sowie Bestandteile des Sicherheitssystems zu nennen. Durch Meßeinrichtungen wird der Strahlungspegel in den Aggregaten und in den Betriebsräumen laufend überwacht. Durch optische und akustische Signale werden eventuell auftretende Erhöhungen des Strahlungspegels sofort angezeigt. Unzulässige Leistungserhöhungen im Reaktor werden automatisch abgebaut, bzw. der Reaktor wird automatisch mit Hilfe von *Havariestäben* in kürzester Zeit stillgelegt. Dem Havariefall trägt auch das *Containment*, eine drucksichere Stahl- bzw. Betonschale Rechnung, die den gesamten Primärkreislauf einschließt. Auf diese Weise werden die im Havariefall unter Druck freigesetzten radioaktiv verseuchten Gase und Flüssigkeiten von der Umgebung ferngehalten.
Die modernen mit thermischen Neutronen betriebenen Kernkraftwerke können aufgrund ihres komplexen Sicherheitssystems und entsprechend den inzwischen gewonnenen Betriebserfahrungen durchaus in dichtbesiedelten Gegenden errichtet werden.
Auf lange Sicht wird es nötig, ein System von Kernkraftwerken durch Anlagen zu ergänzen, die einen optimalen Kreislauf des Kernbrennstoffs garantieren. Dazu gehören Anlagen, in denen Uranium- bzw. Thoriumerze aufbereitet werden, Anlagen zur Herstellung und Anreicherung des Kernbrennstoffs, zur Fertigung von

Brennelementen und Brennstoffkassetten und zum Transport, zur Lagerung sowie zur Wiederaufbereitung von abgebranntem Brennstoff oder bestrahltem Brutstoff. Nur ein koordiniertes Zusammenspiel dieser Anlagen mit den Kernkraftwerken ermöglicht einen effektiven Einsatz der Kernenergetik in großen Maßstäben.

3.5.2.3. Kernkraftwerke. Die Aufgabe eines Kernkraftwerks besteht in der Umwandlung der Kernenergie in mechanische bzw. elektrische Energie. Die wesentlichen Besonderheiten gegenüber einer konventionellen Energieumwandlung sind dabei das Auftreten von radioaktiver Strahlung in den verschiedensten Formen und die reaktorphysikalischen Anforderungen, die zur Aufrechterhaltung einer steuerbaren Kettenreaktion zu erfüllen sind. Die wärmetechnischen Aufgaben, die bei solchen Energieumwandlungen eine wesentliche Rolle spielen, sind konventioneller Art und unterscheiden sich von den entsprechenden bekannten technischen Aufgaben im wesentlichen nur durch die Parameter.

Es gibt entsprechend den verschiedenen Kopplungsmöglichkeiten von Brennstoff, Moderator und Kühlmittel sowie den unterschiedlichsten Arbeits- und Betriebsbedingungen eine große Anzahl verschiedener *Reaktortypen*, die nach den unterschiedlichsten Merkmalen klassifiziert werden können. Vom physikalischen Standpunkt erscheint die Unterscheidung nach der Neutronenenergie, bei der die Kernspaltungen ausgelöst werden und der grundsätzlichen Anordnung von Brennstoff, Moderator und Kühlmittel bedeutungsvoll. Im ersten Fall läßt sich hauptsächlich zwischen *schnellen* und *thermischen* und im zweiten Fall zwischen *heterogenen* und *homogenen* Reaktoren unterscheiden. Da für industrielle Kernkraftwerke bis etwa 1990 nur thermische Reaktoren mit vorwiegend heterogener Anordnung zum Einsatz gelangen, reichen diese Einteilungsprinzipien nicht aus.

Für technische Anwendungen erscheint die Einteilung der Reaktoren nach den Kühlmitteln und deren Eigenschaften am bedeutungsvollsten. Die Praxis hat gezeigt, daß durch das Kühlmittel nicht nur Anreicherungsgrad und Brennstoffart, sondern auch Hüllenmaterial und Geometrie der Spaltzone vorgegeben sind. Damit liegen aber sowohl die Wärmeübertragungsbedingungen im Reaktor als auch die grundsätzliche Anordnung des gesamten Kraftwerks fest. Im Einsatz sind die folgenden Reaktortypen.

3.5.2.3.1. *Druckwasserreaktor.* Er besteht aus einem druckfesten Kessel, der mit Wasser gefüllt ist und in seinem Innern den Brennstoff und die Regeleinrichtung enthält. Als Kühlmittel findet ebenfalls Wasser unter hohem Druck Verwendung. Durch den hohen Druck wird die Wasserverdampfung, die zu Instabilitäten führt, vermieden. Das Kühlwasser gibt im Wärmetauscher die gespeicherte Wärme an das Wasser im Sekundärkreislauf ab, das als Sattdampf mit etwa $5 \cdot 10^6$ Pa Druck auf die Turbine geführt wird. Das Kühlmittel kann im ersten Kreislauf bei Drücken von etwa $16 \cdot 10^6$ Pa nur Temperaturen bis 350 °C erreichen.

3.5.2.3.2. *Siedewasserreaktor.* Er ist dem Druckwasserreaktor ähnlich. Das Kühlmittel wird jedoch z. T. im Reaktor verdampft und auf die Turbine geführt. Allerdings fällt mit dem Wärmetauscher auch ein sehr wirksamer Schutz des Dampfkreislaufs gegen radioaktive Spaltprodukte fort.

Beide Reaktortypen arbeiten in der Regel mit schwach angereichertem Uranium in Metall- bzw. Oxidform und zeichnen sich durch eine hohe Stabilität aus. Bei Schwerwassermoderierung wird Natururanium als Brennstoff benutzt.

3.5.2.3.3. *Gasgekühlter Reaktor.* Er ist ein heterogener graphitmoderierter Reaktor, der mit CO_2-Gas gekühlt und sowohl mit natürlichem Uranium in metallischer Form als auch mit angereichertem Uranium in Oxidform betrieben wird. Während die Kühlmitteltemperatur im ersten Fall auf 670 °C begrenzt ist und dadurch nur ein relativ niedriger Wirkungsgrad des Kraftwerks erreicht wird, werden im zweiten Fall hohe Kühlmitteltemperaturen und somit höhere Wirkungsgrade der Eigenenergieumwandlung erzielt. Vorteilhaft ist, daß durch die Gaskühlung keine Korrosionsprobleme auftreten. Ein Nachteil ist der hohe Energieverbrauch für die Umwälzpumpen des Kühlgases. Außerdem hat der Reaktorkern durch die Graphitmoderierung ein großes Volumen.

3.5.2.3.4. *Schneller Brutreaktor.* Pu-239 ist bei der Verwendung von schnellen Neutronen in der Neutronenökonomie ergiebiger als die beiden anderen Spaltstoffe. Es ist deswegen aussichtsreich, sogenannte schnelle Reaktoren mit diesem Brennstoff zu bauen. Plutonium kommt jedoch in der Natur nicht vor und muß erst aus U-238 durch Neutroneneinfang erbrütet werden (↑ 3.5.1.4.).

Ein schneller Reaktor arbeitet mit einer Anreicherung des Brennstoffs auf 12 bis 30% Spaltmaterial (U-235, Pu-239). Diese hohe Spaltstoffkonzentration und das Fehlen des Moderators führen zu kleinen kritischen Maßen und daher zu ungewöhnlich hohen Leistungsdichten von rund 500 kW pro Liter in der Spaltzone. Es werden Brennstoffelemente aus keramischem Material (Oxide, Carbide, Nitride) und flüssiges Natrium als Kühlmittel verwendet. Bei Plutonium ist vorgesehen, es wegen seines niedrigen Schmelzpunkts in flüssiger Form zu verwenden, wobei es gleich-

zeitig als Wärmeübertragungsmittel dienen soll. Daneben wird mit verschiedenen Plutoniumlegierungen experimentiert.

Die Spaltzone eines schnellen Reaktors ist von einem Brutmantel aus U-238 bzw. Th-232 umgeben, in dem durch Neutroneneinfang neuer Spaltstoff Pu-239 bzw. U-233 erbrütet wird (↑ 3.5.1.4.).

Natrium als Kühlmittel ermöglicht hohe Brutraten und hat gute wärmephysikalische Eigenschaften. Es kann Heißdampf bis 580 °C und $24 \cdot 10^6$ Pa erzeugt werden, so daß sich die für konventionelle Kraftwerke entwickelten Turbinen verwenden lassen. Nicht unerheblich sind beim flüssigen Natrium als Kühlmittel die Korrosionsprobleme.

Die Sicherheits- und Regelungsprobleme sind beim schnellen Reaktor komplizierter als beim thermischen. Die Regelungsreaktion auf Störungen des stationären Betriebs muß außerordentlich schnell erfolgen. Beim natriumgekühlten Betrieb führt ein Kühlmittelverlust zur Reaktivitätserhöhung. Die damit verbundene Wärmeentwicklung ist wegen der hohen Leistungsdichte sehr groß. Durch entsprechende Maßnahmen muß gesichert werden, daß bei Störungen im Kühlsystem die Spaltzone nicht schmilzt und im Falle eines Zusammenschmelzens keine überkritische Anordnung entsteht, die zu einer Kernexplosion führt. Die schnellen Reaktoren haben eine gute Uraniumausnutzung, hohen Abbrand, große Leistung und einen hohen Wirkungsgrad. Nach entsprechenden Entwicklungsarbeiten könnte das zu niedrigen Anlage- und Betriebskosten führen. Wegen der begrenzten Vorkommen des Spaltstoffs U-235 in der Natur ist der schnelle Brüter, der neben Energie neue Spaltstoffe aus dem reichlich vorkommenden U-238 und Th-232 erzeugt, der Reaktor der Zukunft.

4. Elektronen, Lichtquanten und Atome

4.1. Lichtquanten

4.1.1. Messung und Beeinflussung von Strahlungseigenschaften

4.1.1.1. Wellenlänge

4.1.1.1.1. *Spektrometer.* Sie dienen zur Messung der in der elektromagnetischen Strahlung enthaltenen Wellenlängen bzw. Spektralanteile. Die in den verschiedenen Spektralbereichen des elektromagnetischen Spektrums (↑ 2.2.4.3.) verwendeten Spektrometer unterscheiden sich nach Art des wellenlängenselektiven Elements (Tab. 4.1.1-1) und Art des Strahlungsnachweises (↑ 4.1.1.3.).

Die Güte der Spektrometer, d. h. ihre Fähigkeit, zwei um die Wellenlängendifferenz $\Delta\lambda$ sich unterscheidende Spektrallinien noch getrennt zu registrieren, wird quantitativ in den verschiedenen Spektralbereichen durch das *Auflösungsvermögen* $\lambda/\Delta\lambda$ angegeben.

Bei *Prismenspektrometern und -monochromatoren* wird die wellenlängenabhängige Brechung der Strahlung zur Trennung in ihre spektralen Anteile benutzt (Abb. 4.1.1-2). Die erste Linse erzeugt ein paralleles Lichtbündel, die zweite bildet Strahlung gleicher Wellenlänge auf strichförmige Gebiete in der Bildebene ab. Dort kann entweder ein Flächenempfänger, z. B. eine Fotoplatte, eine Fernsehkamera, oder ein Austrittsspalt mit nachfolgendem Strahlungsempfänger (↑ 4.1.1.3.) angeordnet sein. Mit Einprismenspektrometern wird ein Auflösungsvermögen bis etwa 10^4 erreicht. Zur Erhöhung desselben, insbesondere in Spektralbereichen mit geringer Dispersion des Prismenmaterials, werden Spektrometer eingesetzt, in denen die Strahlung das Prisma mehrmals durchläuft.

Tab. 4.1.1-1 Spektrometer für verschiedene Spektralbereiche; λ Wellenlänge

Spektralbereich	λ in nm	wellenlängenselektives Element
Röntgenstrahlung	< 10	Kristallgitter
Vakuumultraviolett	10 bis 200	Konkavgitter
	100	Flußspatprismen
Ultraviolett	200 bis 400	Quarzprismen
sichtbarer Bereich	400 bis 700	Glasprismen
nahes Infrarot	700 bis 1000	Gitter
Infrarot	$1 \cdot 10^3$ bis $5 \cdot 10^3$	Metallhalogenidprismen
	$5 \cdot 10^3$ bis $40 \cdot 10^3$	Gitter
fernes Infrarot	$40 \cdot 10^3$ bis $400 \cdot 10^3$	Echelettegitter
Mikro- und Radiowellen	$> 400 \cdot 10^3$	Mikrowellenelektronik, Klystron

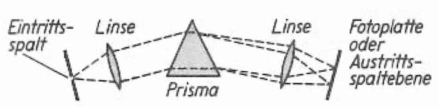


Abb. 4.1.1-2 Strahlengang eines Prismenspektrometers

Bei *Gitterspektrometern* und *-monochromatoren* wird das Prisma durch ein Beugungsgitter (↑ 2.3.2.3.) ersetzt. Man stellt solche Gitter durch Ritzen von Glas- oder Metalloberflächen oder auch fotografisch her. Der Abstand zweier Spalte beträgt bei im sichtbaren oder ultravioletten Spektralbereich einsetzbaren Spektrometern etwa 1 µm. Für das Gebiet des Vakuumultravioletts und der weichen Röntgenstrahlung ($\lambda < 200$ nm) ist die Gitterkonstante $d = 1$ µm zu groß gegen die Wellenlänge. Man läßt deshalb das Licht unter einem sehr kleinen Winkel zur Gitterebene auf das Reflexionsgitter fallen, um den effektiven Furchenabstand zu verringern. Da in diesem Spektralbereich die Luft selbst die Strahlung absorbiert, muß die gesamte Apparatur in einem Vakuumkessel untergebracht sein. Weil im Spektralgebiet $\lambda < 200$ nm keine herkömmlichen Linsen zur Abbildung der Strahlen eingesetzt werden können, werden reflektierende Konkavgitter verwendet. Die Gitterfurchen sind in eine konkave Kugelfläche geritzt. Ein solches Konkavgitter erfüllt zwei Funktionen: die Zerlegung des Lichts und die Abbildung des Spektrums auf die Fotoplatte bzw. die Ebene des Austrittsspalts.

Im mittleren und fernen Infrarotgebiet ist die Gitterkonstante entsprechend größer als im sichtbaren Spektralbereich. Heute werden für das Infrarotgebiet und den sichtbaren Bereich des Spektrums im Gitterspektrometer soge-

nannte *Echelettegitter* eingesetzt. Diese gestatten, infolge der speziellen Formgebung der Gitterfurchen, die gesamte Strahlungsleistung im wesentlichen in eine Beugungsordnung (↑ 2.3.2.3.) zu lenken. Mit Gitterspektrometern kann ein Auflösungsvermögen von 10^5 erreicht werden.

Im Röntgenwellengebiet werden *Kristallgitterspektrometer* verwendet. Ausgenutzt wird dabei die Beugung von Röntgenstrahlen an den mit Atomen besetzten Netzebenen von Kristallen (↑ 6.4.1.).

Bei den *Radiowellenspektrometern* kann im Frequenzgebiet der Radio- und Mikrowellen kohärente monochromatische Strahlung in Schwingkreisen erzeugt werden, die aus einem aktiven Element Energie zugeführt bekommen. Durch Variation von Schwingkreisparametern kann die Frequenz der Strahlung verändert werden. Im Radiowellenspektrometer wird die Leistung am Ausgang der durchstrahlten Probe mit Halbleiterdetektoren in Abhängigkeit von der Frequenz gemessen.

4.1.1.1.2. *Hochauflösende Spektrometer, Interferometer.* Bei hochauflösenden Spektrometern werden zur Wellenlängenselektion verschiedene Interferometer wie Stufengitter, Lummer-Gehrke-Platte und Fabry-Perot-Interferometer eingesetzt, bei denen das hohe Auflösungsvermögen durch einen großen Gangunterschied zwischen den interferierenden Lichtbündeln erreicht wird. Das *Stufengitter* besteht aus vielen übereinandergelegten Glas- oder Quarzplatten gleicher Dicke (Abb. 4.1.1-3). Die die einzelnen Plattenelemente durchlaufenden Strahlen besitzen nach Durchgang durch das Stufengitter

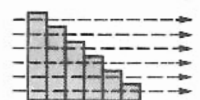


Abb. 4.1.1-3 Stufengitter

zu ihren Nachbarstrahlen einen festen Gangunterschied $b(n - 1)$, wenn b die Dicke und n die Brechzahl eines Plattenelements ist. Die mittels Linse vereinigten Strahlen erzeugen ein Interferenzbild, welches beim Gitter einem Beugungsspektrum hoher Ordnung entspricht.

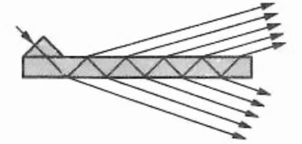


Abb. 4.1.1-4 Lummer-Gehrcke-Platte

Die *Lummer-Gehrke-Platte* besteht aus einer planparallelen Glasplatte hoher Qualität, auf die ein kleines Prisma zur Einkopplung der Strahlung aufgekittet ist (Abb. 4.1.1-4). Die interferierenden Strahlenbündel großen Gangunterschieds werden bei der Lummer-Gehrke-Platte durch Vielfachreflexionen an der Plattenober- und -unterseite erzeugt.

Das *Fabry-Perot-Interferometer* besteht aus zwei ebenen Glas- oder Quarzplatten, deren eine Seite mit einer reflektierenden Schicht bedampft ist. Im Infrarotbereich werden Germaniumplatten verwendet. Die mit der Reflexionsschicht versehenen Plattenseiten werden genau parallel zueinander justiert (Abb. 4.1.1-5). Ähnlich wie bei der Lummer-Gehrke-Platte werden zur Erzielung eines hohen Auflösungsvermögens Interferenzen bei großen Gangunterschieden durch Vielfachreflexionen erzeugt (s. a. 4.1.4.2.). Es ist ein Auflösungsvermögen bis zu 10^8 erreichbar.

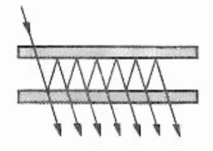


Abb. 4.1.1-5 Fabry-Perot-Interferometer

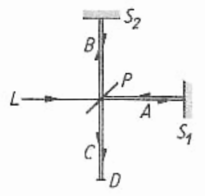


Abb. 4.1.1-6 Michelsoninterferometer

4.1.1.1.3. *Absolutmessung von Wellenlängen, Meterdefinition.* Mit Hilfe von Gitterspektrometern ist es bei genauer Vermessung des Beugungswinkels α und der Gitterkonstanten a möglich, über die Beziehung $a \sin \alpha = m\lambda$ (↑ 2.3.2.3.) Wellenlängen direkt zu bestimmen. Die ersten Absolutmessungen von Lichtwellenlängen mit Beugungsgittern wurden von FRAUNHOFER (1828) und ROWLAND (1887) durchgeführt. Ein Interferenzverfahren zur Absolutmessung von Lichtwellenlängen wurde von MICHELSON (1895) entwickelt. Beim *Michelsoninterferometer* (Abb. 4.1.1-6) wird die von L kommende monochromatische Strahlung mit der halbverspiegelten Teilerplatte P in zwei intensitätsmäßig gleichgroße Anteile aufgeteilt. Ein Teil des Lichts wird von P reflektiert und gelangt über den Lichtweg B zum Spiegel S_2, der durch den Teilerspiegel P hindurchgehende Teil des Lichts läuft über den Lichtweg A zum Spiegel S_1. Beide Spiegel reflektieren die Strahlungsanteile in sich zurück. Nach erneuter Reflexion bzw. erneutem Durchgang durch P kommen beide Teilstrahlen auf dem Lichtweg C zur Interferenz. Je nachdem, ob die Wegdifferenz der Wege B und A ein geradzahliges oder

ungeradzahliges Vielfaches von $\lambda/4$ ist, tritt durch Interferenz bei D Verstärkung oder Auslöschung auf (↑ 2.3.2.1.). Zur Wellenlängenmessung kann der Spiegel S_1 parallel zu sich selbst meßbar verschoben und gleichzeitig der Hell-Dunkel-Wechsel bei D gezählt werden. Die nach diesem Prinzip arbeitenden *Interferenzkomparatoren* gestatten es, die Längendifferenz von zwei *Endmaßen* direkt in Vielfachen und Bruchteilen von λ anzugeben. Mit Interferenzkomparatoren ist es möglich, Längen von etwa 1 m auf 10^{-9} m genau zu vermessen. Lichtwellenlängen konnten damit in bestimmten Bruchteilen des Urmeters angegeben werden. Heute wird umgekehrt die Wellenlänge einer bestimmten Emissionslinie vorgegeben und das Meter als ein Vielfaches derselben definiert. Die zur Meterdefinition ausnutzbaren Emissionslinien müssen durch sehr kleine Linienbreiten $\Delta\nu$ ausgezeichnet sein, damit die Kohärenzlänge $c/\Delta\nu$ (↑ 2.3.2.2.) mindestens in dem Bereich von einigen Dezimetern liegt.

Das Meter wird gegenwärtig als das 165076374-fache der Vakuumwellenlänge einer bestimmten atomaren Emissionslinie des Isotops Krypton-86 definiert (s. a. 4.1.4.7.).

4.1.1.2. Direkte Frequenzmessung

4.1.1.2.1. *Frequenzmessung im Radiowellenbereich.* Im Bereich bis 500 MHz kann mit elektronischen Mitteln eine direkte *Frequenzzählung* vorgenommen werden. Bei Radiowellen mit Frequenzen über 500 MHz wird so verfahren, daß man die Frequenz eines Normalsenders elektronisch vervielfacht und diese vervielfachte Frequenz mit der unbekannten mischt. Gemessen wird dann die Differenzfrequenz zwischen beiden, die dabei im direkt meßbaren Frequenzbereich liegen muß.

4.1.1.2.2. *Lichtmischung.* Durch spezielle Frequenzmischverfahren kann man auch im Infrarot- und sichtbaren Spektralbereich Differenzfrequenzen erzeugen, die mit elektronischen Mitteln unmittelbar meßbar sind (↑ 4.1.1.2.1.). Man geht aus von der Strahlung eines Klystrons, dessen Frequenz mit dem Frequenzstandard, der Caesiumuhr, direkt verglichen wird. Die Frequenz der Klystronstrahlung wird um einen solchen Faktor vervielfacht, daß die entstehende Oberwelle frequenzmäßig mit der Strahlung eines Lasers im fernen Infrarotgebiet verglichen werden kann. Mit Hilfe sogenannter *Punktkontaktdioden* (Metall-Oxid-Metall) wird die elektronisch meßbare Differenzfrequenz der Laserstrahlung und der vervielfachten Klystronstrahlung gebildet. Da Molekülgaslaser bei Verwendung verschiedener Füllgase Emissionslinien besitzen, die über den gesamten infraroten Spektralbereich verteilt sind (↑ 4.1.4.6.), ist es möglich, durch stufenweise Frequenzdifferenzbildung aus dem nahen Infrarotbereich an die Frequenz der Laserstrahlung im fernen Infrarot anzuschließen. Der Anschluß an Frequenzen im sichtbaren Spektralbereich geschieht durch Frequenzvervielfachung der Infrarotlaserstrahlung mit nichtlinear optischen Kristallen (↑ 4.1.4.7.4.).

4.1.1.3. Strahlungsleistung. Die *Strahlungsleistung* ist die von einer Strahlungsquelle pro Zeiteinheit abgegebene Strahlungsenergie. Zu ihrer Messung werden Strahlungsempfänger verwendet. In Tab. 4.1.1-7 sind die in den verschiedenen Spektralbereichen verwendeten Empfänger zusammengestellt. Wichtige Kenngrößen eines Strahlungsdetektors sind seine Zeitkonstante und das Signal-Rausch-Verhältnis der gesamten Nachweiseinrichtung. Die Zeitkonstante gibt die untere Grenze der Zeitdauer von Strahlungsimpulsen bzw. von zeitlichen Perioden bei schwankender Strahlungsleistung an, bis zu der die Strahlungsleistung noch zeitgetreu registriert werden kann. Das Signal-Rausch-Verhältnis bestimmt die untere Nachweisgrenze bei der Strahlungsmessung. In einem sehr weiten Spektralbereich $\lambda = 10$ nm bis 1 µm können Fotoplatten als Strahlungsnachweismittel eingesetzt werden. Die durch die Strahlungseinwirkung hervorgerufene Plattenschwärzung kann mit Fotometern quantitativ

Tab. 4.1.1-7 Strahlungsempfänger für verschiedene Spektralbereiche; λ Wellenlänge

Spektralbereich		Strahlungsempfänger	ausgenutzte Wirkung der Strahlung
Röntgenstrahlung	$\lambda <$ 10 nm	Zählrohr, Szintillationsempfänger	Ionisation von Gasen
Vakuum-UV	$\lambda =$ 10 bis 200 nm		
UV	$\lambda =$ 200 bis 400 nm	Fotoplatte	fotochem. Wirkung
sichtbares Licht	$\lambda =$ 400 bis 800 nm	Vakuumfotozelle, SEV	äußerer Fotoeffekt
nahes IR	$\lambda =$ 700 nm bis 1 µm	Halbleiterempfänger	innerer Fotoeffekt
IR	$\lambda =$ 1 bis 5 µm		
	$\lambda =$ 5 bis 40 µm	Thermoelement, Bolometer	thermische Wirkung
fernes IR	$\lambda =$ 40 bis 400 µm	Golayzelle	

bestimmt werden. Vor der Messung muß mit geeigneten Graufiltern eine Eichkurve für die Fotoplatte aufgenommen werden.

In *thermischen Strahlungsempfängern* wird die Strahlungsenergie in der meist künstlich geschwärzten Empfängerfläche in Wärme umgewandelt. Die resultierende Temperaturerhöhung wird gemessen. Die Empfindlichkeit von thermischen Strahlungsdetektoren ist von der Wellenlänge weitgehend unabhängig. Beim *Bolometer* wird die Temperaturerhöhung der Empfängerfläche über eine Widerstandsänderung elektrisch gemessen.

Thermoelemente bestehen aus drei mit zwei Lötstellen verbundenen Metalldrähten, wobei der mittlere Draht aus einem anderen Metall besteht. Eine der beiden Lötstellen steht mit der Empfängerfläche in thermischem Kontakt und erwärmt sich bei Bestrahlung. Durch die Temperaturdifferenz zwischen den beiden Lötstellen entsteht eine Thermospannung, die gemessen wird. Häufig werden Thermosäulen, das sind mehrere gekoppelte Thermoelemente, verwendet. Die *Golayzelle* ist ein kleiner, gasgefüllter Behälter, dessen eine Behälterwand in Form einer dünnen Membran gestaltet ist. Die durch die einfallende Strahlung hervorgerufene Temperaturerhöhung bewirkt ein Ansteigen des Gasdrucks. Das führt zu einer Auswölbung der Membran. Die Membranverformung wird elektrisch oder mit optischen Verfahren gemessen.

Die Wirkung der *Vakuumfotozellen* beruht auf dem äußeren Fotoeffekt (↑ 1.2.2.2.). Elektronen können nur dann aus der Fotokatode austreten, wenn die auftreffenden Photonen mindestens die Energie der vom Katodenmaterial abhängigen Ablösearbeit der Elektronen besitzen. Die Quantenausbeute gibt die Wahrscheinlichkeit an, mit der ein Photon ein Elektron aus der Fotokatode ablöst. Die Zahl der pro Zeiteinheit ausgelösten Elektronen ist der Strahlungsleistung proportional. Mit Fotozellen ist es möglich, bei hinreichend kleinen Strahlungsleistungen, d. h. bei hinreichend wenig pro Zeitintervall einfallenden Photonen, eine Photonenzählung vorzunehmen. Bei bekannter Quantenausbeute ist die Photonenzählung zurückführbar auf eine Zählung der Fotoelektronen. Moderne Vakuumfotozellen besitzen Stromanstiegszeiten $< 10^{-9}$ s.

Zur Verstärkung des Fotostroms wurden *Sekundärelektronenvervielfacher* (SEV) entwickelt (Abb. 4.1.1-8). Die aus der Fotokatode durch einfallende Photonen ausgelösten Elektronen werden beschleunigt und treffen auf die erste Dynode. Aus dieser und auch aus den nachfolgenden Dynoden werden durch den Elektronenbeschuß Sekundärelektronen ausgelöst. Man erreicht bei genügender Dynodenzahl eine Stromverstärkung um den Faktor 10^8. Sekun-

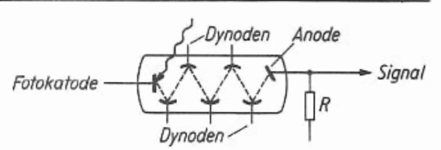


Abb. 4.1.1-8 Sekundärelektronenvervielfacher

därelektronenvervielfacher sind ebenfalls zur Photonenzählung einsetzbar.

Wenn die räumliche Verteilung der einfallenden Strahlung, z. B. das hinter einem Spektrographen erscheinende Spektrum, registriert werden soll, werden *Bildwandlerröhren* verwendet. Die Zahl der aus verschiedenen Punkten der Katodenfläche austretenden Elektronen ist der räumlichen Intensitätsverteilung proportional. Die Elektronen werden beschleunigt und elektronenoptisch auf einen Leuchtschirm abgebildet. Mit sogenannten *Streakkameras* [*streak* engl., Streifen] können zusätzlich neben der räumlichen Verteilung der Strahlung zeitliche Änderungen der Strahlungsintensitäten gemessen werden. Man erreicht dies durch eine zeitlich definierte Ablenkung des Elektronenstrahls. Streakkameras gestatten Zeitauflösungen bis 10^{-12} s. Eine weitere Variante, räumliche Intensitätsverteilungen zu messen, beruht darauf, daß die Fotoelektronen auf einer aus vielen Einzelelementen bestehenden Katodenfläche in Form eines Ladungsbilds gespeichert werden und dieses wie bei einer Fernsehkamera mit einem Elektronenstrahl abgetastet wird.

Die Wirkung von *Halbleiterempfängern* basiert darauf, daß durch Bestrahlung der Halbleiterelemente eine Leitfähigkeitsänderung hervorgerufen wird. Dies kann durch Erzeugung von Ladungsträgern geschehen, die bei Strahlungseinwirkung aus dem Valenz- ins Leitungsband bzw. bei dotierten Halbleitern aus Donatorniveaus ins Leitungsband (↑ 8.3.1.) übergehen. Leitfähigkeitsänderungen durch Ladungsträgererzeugung wird bei Fotoleitungsempfängern, Fotodioden und Fototransistoren ausgenutzt. Man erreicht mit solchen Empfängern Anstiegszeiten von etwa 1 ns. Möglich ist auch eine Leitfähigkeitsänderung dadurch, daß die Strahlung die Ladungsträger innerhalb eines Bands durch sogenannte Intrabandübergänge anregt und ihre Beweglichkeit ändert. Dieser Effekt wird bei der *Gunndiode* ausgenutzt (↑ 8.3.5.5.). Ein anderer Weg, schwache Infrarot-Strahlungsleistungen zu messen, ist der, daß in einem nichtlinear optischen Kristall (↑ 4.1.4.7.4.) die Infrarotstrahlung mit einer sichtbaren Laserstrahlung überlagert wird. Die Strahlungen bei den Mischfrequenzen liegen im sichtbaren Spektralbereich und können sehr empfindlich nachgewiesen werden. Im Spektralbereich der

Röntgenstrahlung werden zur Messung der Strahlungsleistung und -energie Zählrohre oder Szintillationszähler (↑ 3.4.3.3.1.) verwendet.

4.1.1.4. Polarisationszustand. Mit kohärenten Strahlungsquellen, z. B. Radiosendern oder Lasern, kann direkt Strahlung mit definiertem Polarisationszustand, z. B. linear polarisierte oder zirkular polarisierte Strahlung (↑ 2.2.4.2.), erzeugt werden.

Thermische Strahlungsquellen emittieren im allgemeinen Wellenzüge mit vielen verschiedenen Polarisationsrichtungen, so daß ein zeitlich und räumlich mittelnder Strahlungsempfänger keine bevorzugte Polarisationsrichtung registrieren kann. Man spricht dann von natürlicher Strahlung. Atome, die sich in äußeren elektrischen oder magnetischen Feldern befinden oder in Gitter bestimmter Kristallstruktur eingebaut sind, senden teilweise polarisiertes Licht aus.

Es gibt mehrere Möglichkeiten, aus natürlicher Strahlung solche mit einer Polarisationsrichtung auszusondern. Bei *Reflexion* an Oberflächen durchsichtiger isotroper Medien, z. B. Glas, ist der reflektierte Lichtstrahl im allgemeinen teilweise polarisiert. Der Polarisationsgrad des reflektierten Lichts hängt ab vom Einfallswinkel α und der Brechzahl n des Glases. Vollständig linear polarisiert ist der reflektierte Strahl dann, wenn der Einfallswinkel gleich dem durch $\tan \alpha_B = n$ definierten Brewsterwinkel ist. Bei Reflexion an Metalloberflächen entsteht aus linear polarisiertem Licht elliptisch polarisiertes (↑ 2.2.4.2.). Der bei der *Brechung* an isotropen Medien in das Medium gebrochene Lichtstrahl ist teilweise polarisiert.

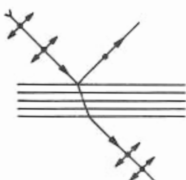


Abb. 4.1.1-9 Erzeugung von linear polarisiertem Licht durch Mehrfachbrechung

Fast vollständig linear polarisiertes Licht entsteht, wenn das Licht viele aufeinander liegende Glasplatten durchläuft (Abb. 4.1.1-9). Strahlt man Licht in optisch anisotrope Kristalle schief zu deren optischer Achse ein, so wird das Licht in einen ordentlichen und einen außerordentlichen Strahl zerlegt, die beide zueinander senkrecht polarisiert sind. Diese Erscheinung wird als *Doppelbrechung* bezeichnet. Die Doppelbrechung wird bei Polarisationsprismen zur Erzeugung von linear polarisiertem Licht ausgenutzt. Bestimmte anisotrope Kristalle oder ausgerichtete Moleküle absorbieren die Strahlung verschiedener Polarisationsrichtungen unterschiedlich stark. Man bezeichnet solche Materialien als dichroitisch. Der *Dichroismus* wird bei Polarisationsfiltern zur Erzeugung von linear polarisiertem Licht ausgenutzt. Das Verhalten dichroitischer Materialien im optischen Bereich ist vergleichbar mit der polarisierenden Wirkung von parallel zueinander aufgespannten Metalldrähten auf unpolarisierte Radiowellen. Von solch einem Drahtgitter werden nur Radiowellen durchgelassen, deren Vektor der elektrischen Feldstärke senkrecht auf der Drahtrichtung steht.

Die Erzeugung von elliptisch bzw. zirkular polarisierter Strahlung (↑ 2.2.4.2.) geschieht im optischen Bereich unter Verwendung anisotroper Kristalle. Mittels Polarisationsprisma oder Polarisationsfilter wird linear polarisiertes Licht hergestellt. Dieses wird in eine Kristallplatte eingestrahlt und in zwei senkrecht zueinander polarisierte Teilstrahlen zerlegt. Beide Strahlen durchlaufen den Kristall mit unterschiedlicher Phasengeschwindigkeit. Ihre Gangunterschiede am Plattenausgang hängen von der Plattendicke ab. Zirkular polarisiertes Licht entsteht, wenn die Intensitäten beider Teilstrahlen gleich sind und der Gangunterschied nach Durchlauf der Kristallplatte $\lambda/4$ beträgt, solche Kristallplatten werden $\lambda/4$-Platten genannt. Zur Erzeugung von elliptisch polarisiertem Licht bestimmter Elliptizität und Achsenlage werden Anordnungen verwendet, die aus zwei keilförmigen Kristallen mit zueinander senkrecht stehenden optischen Achsen bestehen und hinsichtlich ihrer Dicke einstellbar sind. Polarisatoren und $\lambda/4$-Plättchen sowie Kombinationen aus beiden dienen auch zur Untersuchung des Polarisationszustands der Strahlung.

Die von einem streuenden Medium ausgehende Strahlung ist bei Beobachtung senkrecht zur Strahlrichtung fast vollständig linear polarisiert. Der Vektor der elektrischen Feldstärke der Streustrahlung steht senkrecht auf der Ebene, die Strahl- und Beobachtungsrichtung bilden. Bei Einstrahlung von polarisierter Strahlung hängt die Streuintensität entsprechend davon ab, unter welchem Winkel zur Polarisationsrichtung beobachtet wird. Darauf beruht die Methode von BARKLA zur Untersuchung des Polarisationszustands von Röntgenstrahlung.

4.1.2. Strahlungsgesetze

4.1.2.1. Wärmestrahlung

4.1.2.1.1. *Lambertsches und Kirchhoffsches Gesetz.* Die thermische Anregung von Atomen, Molekülen und Festkörpern spielt für die Lichterzeugung eine wichtige Rolle. Die thermisch angeregte Strahlung wird als Wärmestrahlung oder Temperaturstrahlung bezeichnet. Die ex-

perimentellen Erfahrungen über die *Wärmestrahlung* besagen:
Alle Körper strahlen sich gegenseitig Energie zu, so daß die wärmeren abgekühlt, die kälteren erwärmt werden. Die von einem Körper abgestrahlte Leistung wächst mit seiner Temperatur an und ist unabhängig von der Umgebung.
Mit der Temperatur verändert sich die spektrale Zusammensetzung der Strahlung.
Bei gleicher Temperatur strahlt ein lichtabsorbierender Körper stärker als ein lichtdurchlässiger.
Das *Lambertsche Gesetz* gibt an, wie die von einem erhitzten Flächenelement Δa_1 einem anderen Flächenelement Δa_2 zugestrahlte Leistung ΔP von der Geometrie der Anordnung abhängt; es gilt $\Delta P \sim \Delta a_1 \cos \vartheta \Delta \Omega$, wobei $\Delta \Omega = a_2/R^2$ der Raumwinkel ist, unter dem der Empfänger vom Sender aus gesehen wird (Abb. 4.1.2-1). Die Proportionalität der Strahlungsleistung zu Δa_1 und $\Delta \Omega$ folgt aus einfachen

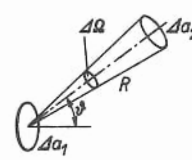


Abb. 4.1.2-1 Relative Lage von strahlender Fläche Δa_1 und empfangender Fläche Δa_2

geometrischen Überlegungen. Die Proportionalität der Strahlungsleistung zu $\cos \vartheta$ kann dagegen nur aus dem Experiment gewonnen werden; streng gilt diese Beziehung nur für die Hohlraumstrahlung (↑ 4.1.2.1.2.). Für einen Lambertstrahler fester Temperatur ist die Strahlungsdichte $s = \dfrac{\Delta P/\cos \vartheta}{\Delta \Omega \Delta a_1}$ eine Konstante.
Das *Kirchhoffsche Gesetz* vergleicht die Strahlungsdichte s verschiedener, bei gleicher Temperatur strahlender Körper; es gilt $s_1/A_1 = s_2/A_2 = s_i/A_i$, wobei als A_i das Absorptionsvermögen eines Körpers i durch den Quotienten aus absorbierter und einfallender Strahlungsleistung definiert ist. Das Kirchhoffsche Gesetz folgt aus allgemeinen thermodynamischen Gesetzen und kann z. B. aus der Überlegung hergeleitet werden, daß im Gleichgewicht zwischen zwei gleichgroßen Körpern 1 und 2, deren Abstand klein gegen ihre Abmessungen ist, der Körper 1 dem Körper 2 gerade so viel zustrahlen muß, wie er von dessen Strahlung absorbiert. Bei spektraler Zerlegung der Strahlung gilt das Kirchhoffsche Gesetz für jedes Spektralintervall und damit auch für die Strahlungsdichte pro Frequenzeinheit $s_\nu = \Delta s/\Delta \nu$.
Der *schwarze Körper* besitzt das Absorptionsvermögen 1 und absorbiert demzufolge alle auffallende Strahlung. Man erhält die Strahlungsdichte eines beliebigen Körpers, indem man gemäß dem Kirchhoffschen Gesetz die Strahlungsdichte des schwarzen Körpers s gleicher Temperatur mit dem Absorptionsvermögen A_i multipliziert $s_i = sA_i$. Die Strahlungsdichte s des schwarzen Körpers kann also nur von der Temperatur abhängen, die Strahlungsdichte pro Frequenzeinheit s_ν nur von Temperatur und Frequenz; s ist also eine universelle Funktion von ν und T. Seit Ende des vorigen Jahrhunderts wurden sehr große experimentelle und theoretische Anstrengungen gemacht, um $s_\nu(\nu,T)$ zu bestimmen (s. a. 4.1.2.1.3.).

4.1.2.1.2. *Hohlraumstrahlung.* Der Grenzfall des schwarzen Körpers läßt sich durch ein kleines Loch in den erhitzten Wänden eines großen Hohlraums annähern, weil alle auf das Loch auftreffende Strahlung im Inneren sehr oft reflektiert und bei den mehrfachen Umläufen selbst bei kleinem Absorptionsvermögen nahezu vollständig absorbiert wird. Sehr genaue Messungen der spektralen Zusammensetzung der Hohlraumstrahlung und ihrer Abhängigkeit von der Temperatur wurden um die Jahrhundertwende von LUMMER und PRINGSHEIM, später von PASCHEN, KURLBAUM und RUBENS vorgenommen.
In einem geschlossenen Hohlraum mit einer konstanten Wandtemperatur T bildet sich im Inneren eine *elektromagnetische Gleichgewichtsstrahlung* heraus, bei der die von den Wänden abgestrahlte Leistung im Mittel gleich der von ihnen absorbierten ist. Die spektrale Energiedichte im Hohlraum, d. h. die Energie pro Volumenintervall und Frequenz $E_{V,\nu}(\nu,T) = \Delta E/(V\Delta \nu)$, ist eine universelle Funktion von Frequenz und Temperatur. Die spektrale Strahlungsdichte $s_\nu(\nu,T)$ der Hohlraumstrahlung, die durch Auskopplung durch ein kleines Loch im Hohlraum gewonnen wird, hängt über $s_\nu(\nu,T) = \dfrac{c}{4\pi} E_{V,\nu}$ unmittelbar mit der spektralen Energiedichte im Hohlraum zusammen.
Hohlraummoden sind Eigenschwingungen (s. a. 2.3.3.) des elektromagnetischen Felds innerhalb eines *Hohlraumresonators*, nach denen jede Feldverteilung zerlegt werden kann.
Die Zahl der an einem physikalischen Vorgang beteiligten Hohlraummoden ist im optischen Gebiet im allgemeinen sehr groß, weil die Abmessungen der Resonatoren sehr groß gegen die Wellenlänge sind. Die Gesamtzahl der Hohlraummoden $\Delta N(\nu)$ im Frequenzintervall $\Delta \nu$ um die Frequenz ν kann als Produkt aus einer Modendichte $N_{V,\nu}$, d. h. einer Modenzahl pro Volumen- und Frequenzeinheit, mit dem Hohlraumvolumen $V = L^3$ und dem Frequenzintervall $\Delta \nu$ dargestellt werden:

(1) $\quad \Delta N(\nu) = N_{V,\nu} V \Delta \nu = (8\pi\nu^2/c^3)\, V\Delta \nu.$

Die Modendichte besitzt den aus dieser Gleichung ablesbaren Wert, wobei bereits ein Faktor 2 für die beiden unabhängigen Polarisa-

tionen jeder einzelnen stehenden Welle berücksichtigt ist. Die spektrale Energiedichte

(2) $E_{V,\nu} = N_{V,\nu}\varepsilon(\nu, T)$

ergibt sich aus der Modendichte und der mittleren Energie einer einzelnen Mode, die als ein Strahlungsoszillator aufgefaßt werden kann. Nach der klassischen Statistik (↑ 2.4.4.) hat die mittlere Energie $\bar{\varepsilon}$ eines harmonischen Oszillators, der *kontinuierlich* alle Energiewerte ε zwischen 0 und ∞ annehmen kann, den Wert $\bar{\varepsilon} = k_B T$, wobei k_B die Boltzmannkonstante und T die absolute Temperatur bedeuten. Dies führt zu dem nach RAYLEIGH und JEANS benannten Strahlungsgesetz, das mit dem Experiment nur für kleine Frequenzen übereinstimmt, d. h., das Konzept der klassischen elektromagnetischen Wellentheorie versagt für hohe Frequenzen.

4.1.2.1.3. *Plancksches Strahlungsgesetz.* Um Übereinstimmung zwischen Theorie und Experiment für alle Frequenzen zu erzielen, mußte PLANCK einen radikalen Bruch mit den klassischen Vorstellungen vollziehen (s. a. 1.2. und 2.5.). Er postulierte – im völligen Gegensatz zur klassischen Physik – ein diskontinuierliches Energiespektrum der Oszillatoren, die den Hohlraummoden zugeordnet sind; er „quantisierte" die Oszillatoren. Dabei sind die Abstände zwischen den diskreten Energiewerten, den Energieniveaus, der Frequenz des Oszillators proportional $\varepsilon_n = h\nu(n + 1/2)$ mit $n = 0, 1, 2, \ldots$; die Proportionalitätskonstante h wird als *Plancksches Wirkungsquantum* (↑ 1.2.5.) bezeichnet. Mit den diskreten Energieniveaus ε_n kann man gemäß dem Boltzmannschen Verteilungsgesetz der Statistik die mittlere Energie der quantisierten Strahlungsoszillatoren berechnen: $\bar{\varepsilon} = h\nu/(e^{h\nu/k_B T} - 1)$. Diese Beziehung geht nur für kleine Frequenzen oder lange Wellen ($h\nu \ll k_B T$) in das klassische Resultat $\bar{\varepsilon} = k_B T$ über. Für

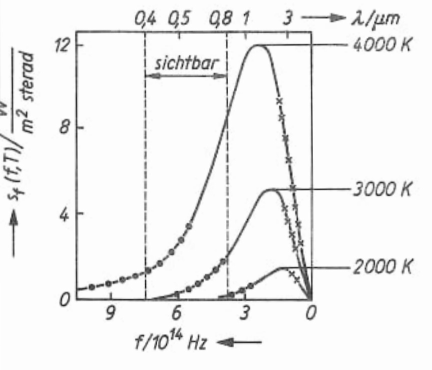


Abb. 4.1.2-2 Plancksches Strahlungsgesetz; mit *Punkten* ist der Gültigkeitsbereich des Wienschen, mit *Kreuzen* der des Rayleigh-Jeansschen Strahlungsgesetzes eingetragen

die Energiedichte des elektromagnetischen Felds und damit auch die spektrale Strahlungsdichte des schwarzen Körpers $s_\nu(\nu,T)$ folgt das *Plancksche Strahlungsgesetz* (Abb. 4.1.2-2):

$$(3) \quad E_{V,\nu}(\nu, T) = \frac{4\pi}{c} s_\nu(\nu, T) = \frac{8\pi\nu^2 h\nu\Delta\nu}{(e^{h\nu/k_B T} - 1)c^3}.$$

Aus diesem folgt speziell das *Wiensche Verschiebungsgesetz* $\lambda_{max}T = \text{const.}$ (const. $\approx 0{,}2\, hc/k_B$), d. h. die Wellenlänge λ_{max}, für die die spektrale Energiedichte ihr Maximum erreicht, ist der absoluten Temperatur umgekehrt proportional. Außerdem folgt das *Stefan-Boltzmannsche Gesetz* für die Gesamtemission eines schwarzen Körpers in den Halbraum:

$$(4) \quad s(T) = \sigma T^4 \quad \text{mit} \quad \sigma = 2\pi^5 k_B^4/15\, c^2 h^3 \approx 5{,}67 \cdot 10^{-8}\ \text{Wm}^{-2}\text{K}^{-4}.$$

4.1.2.2. Lichtquanten. Mit den Arbeiten von PLANCK wurde die Quantisierung des Strahlungsfelds vollzogen. Danach kann die Energie einer Eigenschwingung des elektromagnetischen Felds, d. h. einer Mode, nur bestimmte diskrete Werte annehmen. Damit kann sich die Energie der Mode auch nur quantenhaft ändern, durch Aufnahme oder Abgabe von Energieportionen der Göße $h\nu$.

Diese Quantisierung ist nicht auf den Fall der thermischen Gleichgewichtsstrahlung beschränkt, für den sie zunächst betrachtet wurde, sondern gilt für jede beliebige Anregung der Hohlraummoden. Auch die Durchführung der Quantisierung in einem bestimmten Hohlraum stellt keine Einschränkung dar; seine Form und seine Abmessungen können beliebig gewählt werden, und man kann in der Grenze auch zu unendlich großen Ausdehnungen, also zum unendlich freien Raum, übergehen. Auch den fortlaufenden Wellen im freien Raum sind harmonische Oszillatoren äquivalent, weshalb man sie in gleicher Weise quantisieren kann.

Die Quantisierung einer Mode, d. h. einer Feldschwingung, mit der Frequenz ν kann in zweierlei Weise interpretiert werden. Einmal kann man die Anregung dieser Mode mit der Energie $nh\nu$ als n-te Anregungsstufe eines harmonischen Oszillators mit diskreten Energiestufen $\varepsilon_n = h\nu(n + 1/2)$ verstehen. Die Erhöhung bzw. Erniedrigung der Energie der Mode – durch Emission und Einkopplung von Energie bzw. Absorption und Auskopplung von Energie – wird durch Übergänge zwischen den Niveaus der harmonischen Oszillatoren beschrieben. Zum anderen kann man die Anregung dieser Mode mit der Energie $nh\nu$ durch n wechselwirkungsfreie „Teilchen" der Energie $h\nu$ beschreiben; diese „Teilchen" werden als Lichtquanten bezeichnet. Eine Interpretation der Quantisierung des Strahlungsfelds mit Hilfe von Lichtquanten oder Photonen wurde zum ersten Mal 1905 von EINSTEIN bei der Behand-

lung des äußeren lichtelektrischen Effekts (↑ 1.2.2.2.) vorgenommen.

Die Lichtquanten sind Elementaranregungen der *Eigenschwingungen* oder *Moden* des elektromagnetischen Felds (↑ 2.3.3.). Das Lichtquant wird durch die *Frequenz*, die *Ausbreitungsrichtung* und die *Polarisation* der Mode des Strahlungsfelds charakterisiert. Ein Photon in einer Eigenschwingung mit der Frequenz ν besitzt die Energie $h\nu$, den Impulsbetrag $h\nu/c$ und die Ruhmasse Null (↑ 2.1.3.). Das Lichtquant bewegt sich in allen Koordinatensystemen mit Lichtgeschwindigkeit.

Während alle bekannten Erscheinungen der Lichtausbreitung im freien Raum mit der klassischen Theorie elektromagnetischer Felder von MAXWELL (↑ 2.2.4.) richtig beschrieben werden, muß die Lichtquantenvorstellung zur Erklärung der Wechselwirkung zwischen Licht und Materie herangezogen werden. Die dabei ablaufenden Grundprozesse der Erzeugung und Vernichtung von Lichtquanten – entsprechend einem Emissions- oder Absorptionsvorgang – sowie die Umwandlung von einem Lichtquant in ein solches zu einer anderen Mode – entsprechend einer Lichtstreuung – können durch die Wechselwirkung des elektromagnetischen Felds mit freien oder gebundenen geladenen Teilchen bewirkt werden (↑ 4.1.3.).

4.1.3. Wechselwirkung von Strahlung mit atomaren Systemen

4.1.3.1. Absorption und Emission. Durch Absorption eines Photons der Energie $h\nu = E_2 - E_1$ kann ein Atom, welches sich im Zustand mit der Energie E_1 befindet, in den angeregten Zustand mit der Energie E_2 überführt werden. Umgekehrt wird bei dem Übergang $2 \rightarrow 1$ ein Photon der Energie $h\nu$ emittiert. Absorptions- und Emissionsprozesse haben zufälligen Charakter; der Zeitpunkt eines Absorptions- bzw. Emissionsvorgangs kann nicht vorhergesagt werden. Es kann nur die Wahrscheinlichkeit dafür angegeben werden, daß ein Atom innerhalb eines bestimmten Zeitintervalls ein Photon absorbiert bzw. emittiert. Die Wahrscheinlichkeit pro Zeiteinheit $\left(\frac{\mathrm{d}W}{\mathrm{d}t}\right)_l^A$ (↑ 2.4.3.4.) für die Absorption eines Photons aus einer Mode l (↑ 4.1.2.1.) mit der Frequenz ν, ist der Photonenzahl n_l in dieser Mode proportional; die entsprechende Übergangswahrscheinlichkeit $\left(\frac{\mathrm{d}W}{\mathrm{d}t}\right)_l^E$ für die Emission eines Photons in die Mode l durch das angeregte Atom ist proportional zu $(n_l + 1)$. Die Übergangswahrscheinlichkeit für Emission läßt sich demgemäß in zwei Anteile aufspalten. Der erste Anteil ist der Photonenzahl in der Mode l vor dem Emissionsakt proportional. Man bezeichnet diese Emission, weil sie durch das äußere Strahlungsfeld bewirkt wird, als *stimulierte Emission* $\left(\frac{\mathrm{d}W}{\mathrm{d}t}\right)_l^{\mathrm{stim}\,E} \sim n_l$. Der zweite Anteil ist von den Photonenzahlen in den Moden des eingestrahlten elektromagnetischen Felds unabhängig. Er ist auch dann vorhanden, wenn das Atom keinem äußeren Strahlungsfeld ausgesetzt ist. Man bezeichnet deshalb diesen Emissionsvorgang als *spontane Emission*. Die Wahrscheinlichkeit pro Zeiteinheit, mit der vom angeregten Atom überhaupt ein Photon der Energie $h\nu$ spontan emittiert wird, erhält man durch Summation der Übergangswahrscheinlichkeiten für spontane Emission in die verschiedenen Moden mit der Frequenz ν, die sich hinsichtlich Polarisations- und Ausbreitungsrichtung unterscheiden. Die Gesamtwahrscheinlichkeit für einen Übergang des Atoms aus dem Anfangszustand 2 in den Endzustand 1 durch spontane Emission wird als Einsteinkoeffizient für spontane Emission A_{12} bezeichnet. Während bei der spontanen Emission ein Photon in jede Mode mit der Frequenz ν emittiert werden kann, kann ein Photon bei der stimulierten Emission nur in die Moden mit der Frequenz ν emittiert werden, in denen schon vor der Emission Photonen vorhanden waren. Es wird bei der stimulierten Emission Strahlung ausgesandt, die sich hinsichtlich Frequenz, Ausbreitungs- und Polarisationsrichtung nicht von der einfallenden Strahlung unterscheidet. Man spricht davon, daß das eingestrahlte Licht durch stimulierte Emissionsvorgänge verstärkt wird (s. a. 4.1.4.).

Die Wahrscheinlichkeit für einen Atomübergang durch Absorption irgend eines Photons aus den Moden mit der Frequenz ν bzw. durch stimulierte Emission eines Photons in irgend eine dieser Moden erhält man durch Addition aller $\left(\frac{\mathrm{d}W}{\mathrm{d}t}\right)_l^A$ bzw. $\left(\frac{\mathrm{d}W}{\mathrm{d}t}\right)_l^{\mathrm{stim}\,E}$. Es ergibt sich $\left(\frac{\mathrm{d}W}{\mathrm{d}t}\right)^A = C_{21}n$ und $\left(\frac{\mathrm{d}W}{\mathrm{d}t}\right)^{\mathrm{stim}\,E} = C_{12}n$. Bei nichtentarteten Atomniveaus (↑ 4.2.) ist $C_{12} = C_{21}$. Für den Fall, daß eine isotrope und unpolarisierte Strahlung auf das Atom einwirkt, ist der Proportionalitätsfaktor C_{12} gleich dem Einsteinkoeffizienten für spontane Emission A_{12}. n ist die mittlere Photonenzahl pro Mode bei der Frequenz ν. Für isotrope und unpolarisierte Strahlung hängt n in einfacher Weise mit der spektralen Energiedichte $\varrho_\nu = \Delta E/V\Delta\nu$ (s. a. 4.1.2.1.2.) des Strahlungsfelds zusammen $n = c^3\varrho_\nu/8\pi h\nu^3$. Die Übergangswahrscheinlichkeiten für Absorption und stimulierte Emission sind also der spektralen Energiedichte des Strahlungsfelds proportional, wobei die Proportionalitätsfaktoren B_{21} bzw. B_{12} als Einsteinkoeffizienten für Absorption bzw. sti-

mulierte Emission bezeichnet werden. Nach obigen Formeln besteht zwischen dem Einsteinkoeffizienten für spontane Emission A_{12} und dem für stimulierte Emission die Beziehung $A_{12} = 8\pi h\nu^3 B_{12}/c^3$.

4.1.3.2. Linienbreiten und Lebensdauern. In den vorangegangenen Abschnitten wurden die Energieniveaus des Atoms als unendlich schmal angenommen. Unendlich schmale Energieniveaus und damit verbunden unendlich scharfe Spektrallinien existieren in der Realität nicht.

Die endlichen Breiten der Energieniveaus sind eine Folge der Wechselwirkung des Atoms mit seiner Umgebung. In Festkörpern und in Flüssigkeiten werden infolge der starken Wechselwirkung mit den Nachbaratomen im allgemeinen sehr breite Spektrallinien beobachtet. Eine Aufklärung der Wechselwirkungsmechanismen in Festkörpern und Flüssigkeiten ist in den meisten Fällen äußerst kompliziert. In Gasen führen die Stöße der Gasatome untereinander zur druckabhängigen Druck- oder Stoßverbreiterung (↑ 4.1.3.2.2.). Wird der Gasdruck erniedrigt, geht die Linienbreite des von den Atomen emittierten Lichts gegen einen vom Druck unabhängigen Wert. Die bei sehr verdünnten Gasen beobachtete Linienbreite wird in der Hauptsache durch den *Dopplereffekt* hervorgerufen (↑ 2.3.1.9., 4.1.3.2.3.). Die Dopplerverbreiterung ist eine Folge der verschiedenen Geschwindigkeiten der emittierenden Atome. Aber selbst wenn man ein Experiment durchführt, bei dem nur wechselwirkungsfreie Atome mit einheitlicher Geschwindigkeit – z. B. nur ruhende Atome – zum Meßergebnis beitragen, beobachtet man eine endliche Linienbreite des emittierten Lichts. Sie stellt die Mindestlinienbreite von Spektrallinien dar und wird als natürliche Linienbreite (↑ 4.1.3.2.1.) bezeichnet. Die Mindestlinienbreite wird durch den Prozeß der spontanen Emission selbst bestimmt.

4.1.3.2.1. *Natürliche Linienbreite und Strahlungslebensdauer.* Die spontane Emission (↑ 4.1.3.1.) begrenzt die Verweilzeit eines Atoms im angeregten Zustand. Beim Übergang in den energetisch scharfen Grundzustand hängt die natürliche Linienbreite $\Delta\nu_{\text{nat}} = (2\pi\tau)^{-1}$ unmittelbar mit der durch den spontanen Emissionsprozeß gegebenen mittleren Lebensdauer, der *Strahlungslebensdauer* τ des angeregten Atoms zusammen. Die Lebensdauer ist gleich dem Kehrwert der Einsteinschen Übergangswahrscheinlichkeit A_{12} für spontane Emission. Im klassischen Bild begrenzt die *Strahlungsdämpfung* die Zeitdauer τ des Emissionsvorgangs und damit die Länge $c\tau$ der emittierten Wellenzüge. Die spektrale Breite des emittierten Lichts ist dann durch die obige Beziehung gegeben. Durch Messungen der Strahlungslebensdauer in sehr verdünnten Gasen kann die natürliche Linienbreite ermittelt werden. Die Dichte der Gase muß dabei so niedrig sein, daß nicht andere Prozesse neben der spontanen Emission die Lebensdauer der angeregten Gasatome begrenzen. Die Lebensdauer kann aus dem exponentiellen Abklingen (↑ 2.4.3.4.) der Fluoreszenzstrahlungsleistung bestimmt werden. Dabei werden die Gasatome durch kurze Lichtimpulse eines Impulslasers (↑ 4.1.4.5.) angeregt. Die Länge der Impulse muß kurz gegen τ sein. Für starke Elektronenübergänge ergeben Messungen und quantentheoretische Rechnungen übereinstimmende Werte für A_{12} von etwa $10^8\,\text{s}^{-1}$. Wegen $\Delta\nu_{\text{nat}} = \dfrac{A_{12}}{2\pi}$ kann die natürliche Linienbreite für solche Übergänge zu $10^7\,\text{s}^{-1}$ bis $10^8\,\text{s}^{-1}$ abgeschätzt werden. Natürlich verbreiterte Spektrallinien besitzen ein sogenanntes Lorentzprofil (Abb. 4.1.3-1):

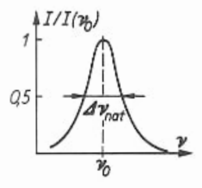


Abb. 4.1.3-1 Lorentzprofil einer Spektrallinie

$I(\nu)/I(\nu_0) = \frac{1}{4}(\Delta\nu_{\text{nat}})^2/[(\nu - \nu_0)^2 + \frac{1}{4}(\Delta\nu_{\text{nat}})^2]$.

$I(\nu_0)$ ist die Intensität der emittierten Strahlung bei der Mittelfrequenz ν_0. Die natürliche Linienbreite $\Delta\nu_{\text{nat}}$ ist gleich der Halbwertsbreite der lorentzförmigen Spektrallinie (↑ 2.1.6.1.).

In entsprechender Weise wird das Atom auch Strahlung in einem bestimmten Gebiet um die Übergangsfrequenz ν_{21} absorbieren. Dieser Sachverhalt kann quantitativ durch Einführung eines frequenzabhängigen Einsteinschen Übergangskoeffizienten für Absorption $B_{21}(\nu) = B_{21}g(\nu - \nu_{21})$ beschrieben werden, wobei das Integral über die Linienformfunktion $g(\nu - \nu_{21})$ auf 1 normiert ist.

4.1.3.2.2. *Stoßverbreiterung und strahlungslose Desaktivierung.* Neben der spontanen Emission existieren weitere Prozesse, die die Lebensdauer der angeregten Atome begrenzen. Bei diesen gehen die Atome *strahlungslos*, also ohne Aussendung von Photonen in energetisch tiefere Zustände über. Strahlungslose Desaktivierungsvorgänge werden als *Relaxationsprozesse* und die durch Relaxationsprozesse bestimmte mittlere Lebensdauer angeregter Atome als *Relaxationszeit* bezeichnet.

In großen Molekülen kann die Anregungsenergie sehr schnell auf viele Freiheitsgrade im Molekül verteilt werden. Relaxationsvorgänge laufen in großen Molekülen in der Regel in Zeiten von 10^{-12} bis 10^{-9} s ab. In Gasen, die aus Atomen oder kleinen Molekülen bestehen, findet eine strahlungslose Desaktivierung angeregter Atome oder Moleküle durch Stöße statt. Ty-

pische Relaxationszeiten in Gasen bei Normalbedingungen liegen im Bereich von 10^{-10} bis 10^{-8} s. Wie die spontane Emission führen auch die anderen lebensdauerbegrenzenden Relaxationsvorgänge zu einer Verbreiterung der Energieniveaus bzw. Spektrallinien. Die spektrale Breite der ursprünglich angeregten Niveaus ist dabei $\Delta\nu_R = (2\pi T_R)^{-1}$. T_R ist die Relaxationszeit. Die in Gasen beobachtete *Druck- oder Stoßverbreiterung* ist aber vorwiegend nicht auf desaktivierende Stöße zwischen den Atomen, sondern auf *phasenverändernde* (elastische) *Stöße* zurückzuführen. Während der Dauer des Emissionsvorgangs von etwa 10^{-8} s (↑ 4.1.3.2.2.) kann z. B. ein Heliumatom unter Normalbedingungen etwa 100mal gestoßen werden. Bei

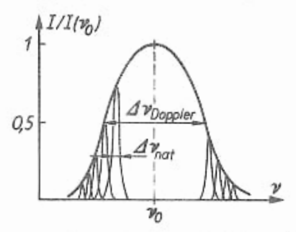


Abb. 4.1.3-2 Dopplerlinie; die *Kurven* unter der Dopplerlinie stellen die natürlich verbreiterten Linien der Gasatome dar

jedem Stoß wird die Phase des emittierten Lichts statistisch verändert; der ursprüngliche Wellenzug wird in kürzere Wellenzüge zerhackt, die eine größere spektrale Breite der Strahlung bedingen (↑ 2.3.2.2. und 2.1.6.1.). Die Halbwertsbreite der stoßverbreiterten Spektrallinie ist angenähert gegeben durch $\Delta\nu_{\text{Stoß}} \approx$ Zahl der Stöße pro Sekunde/π. Da die Stoßzahl dem Gasdruck proportional ist, steigt $\Delta\nu_{\text{Stoß}}$ linear mit wachsendem Druck. In vielen Gasen ist unter Normalbedingungen die Halbwertsbreite der stoßverbreiterten Linien um ein mehrfaches größer als die natürliche Linienbreite. Stoßverbreiterte Spektrallinien haben in der Regel lorentzförmiges Profil.

4.1.3.2.3. *Dopplerverbreiterung.* Ein weiterer Verbreiterungseffekt der Spektrallinien wird durch die Eigenbewegung der strahlenden Atome hervorgerufen. Er beruht auf dem Dopplereffekt (↑ 2.3.1.9.), die daraus resultierende Verbreiterung wird als Dopplerverbreiterung bezeichnet.

In einem Gas soll eine große Zahl von Atomen zum Leuchten angeregt worden sein. Die Übergangsfrequenz für den entsprechenden Atomübergang sei ν_0. Hat nun ein emittierendes Atom in Beobachtungsrichtung, d. h. in Richtung auf den Spektrographen zu, eine Geschwindigkeitskomponente $v_\parallel$ ungleich Null, so empfängt der Spektrograph nicht Licht mit der Frequenz ν_0, sondern wegen des Dopplereffekts gilt $\nu = \nu_0(1 + v_\parallel/c)$. Da in dem Gas während des Beobachtungszeitraums sehr viele Atome mit unterschiedlichen Bewegungsrichtungen und unterschiedlichen Geschwindigkeiten emittieren können, wird mit dem Spektrographen eine verbreiterte Linie gemessen. Das Linienprofil kann unter Benutzung der Maxwellschen Geschwindigkeitsverteilung (↑ 2.4.4.6.) für die Gasatome berechnet werden. Man erhält eine Gaußverteilung (↑ 2.4.3.2.) mit der Halbwertsbreite $\Delta\nu_{\text{Doppler}} = \nu_0[8 \ln (2k_BT/Mc^2)]^{1/2}$, wobei M die Masse der strahlenden Atome ist. Bei einer Temperatur von 300 K ist im sichtbaren Spektralbereich die Dopplerbreite etwa 100 bis 1000mal so groß wie die natürliche Linienbreite (Abb. 4.1.3-2).

4.1.3.3. Mehrquantenprozesse. Bei einem Einquantenprozeß wird vom Atom ein Photon mit der Frequenz $\nu = (E_2 - E_1)/h$ einer der Moden absorbiert bzw. ein Photon in eine Mode emittiert. Von Mehrquantenabsorptions- bzw. Mehrquantenemissionsprozessen spricht man dann, wenn das Atom aus einer oder verschiedenen Moden des Strahlungsfelds *gleichzeitig* mehrere Photonen absorbiert oder in eine oder verschiedene Moden gleichzeitig mehrere Photonen emittiert.

Streuprozesse sind dadurch gekennzeichnet, daß das Atom aus einer Mode des Strahlungsfelds ein Photon absorbiert und gleichzeitig ein Photon in eine andere Mode mit gleicher oder veränderter Frequenz emittiert.

4.1.3.3.1. *Mehrquantenabsorption und -emission.* Mehrphotonenabsorption ist dann möglich, wenn die Summe der Energien mehrerer Photonen gerade der Energiedifferenz zwischen zwei Atomzuständen entspricht (Abb. 4.1.3-3). Zum Nachweis der Zweiphotonenabsorption wird gleichzeitig intensive Laserstrahlung der Frequenz ν_1 und eine vergleichsweise schwache Strahlung mit der Frequenz ν_2 in eine Probe

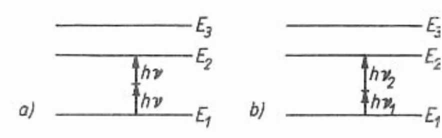


Abb. 4.1.3-3 Zweiphotonenabsorption a) von zwei Photonen gleicher, b) von zwei Photonen verschiedener Frequenz

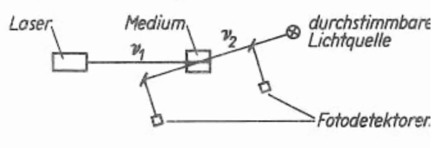


Abb. 4.1.3-4 Nachweis der Zweiphotonenabsorption

eingestrahlt. Mittels Strahlungsempfängern vor und hinter der Probe wird die Transmission der Strahlung bei der Frequenz ν_2 gemessen (Abb. 4.1.3-4). Ist das Atom zum Anfangszeitpunkt im angeregten Zustand der Energie E_2, kann es durch spontane Mehrphotonenemission in den Zustand E_1 übergehen. Die Summe der Frequenzen der emittierten Photonen ist gleich $(E_2 - E_1)/h$.

Experimentell konnte die Zweiphotonenemission dadurch nachgewiesen werden, daß in ein Medium, dessen Atome sich vor der Bestrahlung im angeregten Zustand 2 befanden, Laserstrahlung der Frequenz ν_1 eingestrahlt wurde. Mittels empfindlicher Strahlungsempfänger konnte dann zusätzlich eine aus dem Medium austretende Strahlung der Frequenz $\nu_2 = \frac{E_2 - E_1}{h} - \nu_1$ beobachtet werden.

4.1.3.4.2. *Rayleigh-, Raman- und Brillouinstreuung.* Die Rayleighstreuung ist dadurch gekennzeichnet, daß bei nichtresonanter Einstrahlung einer elektromagnetischen Strahlung in ein Medium, Photonen der Frequenz ν absorbiert und *gleichzeitig* solche der gleichen Frequenz emittiert werden. Die klassische Physik gestattet eine Erklärung der Rayleighstreuung auf folgende Weise: Durch die einwirkende elektromagnetische Welle werden die Atome des Mediums polarisiert. Das induzierte Dipolmoment schwingt mit der Frequenz im Takt mit der einfallenden Welle ν. Jedes Atom sendet seinerseits wieder elektromagnetische Wellen der Frequenz ν aus.

Bei der Ramanstreuung hat das emittierte Photon eine andere Frequenz als das absorbierte. Die Frequenz ν_2 des emittierten Photons kann kleiner (Stokesstreuung) oder größer (Antistokesstreuung) sein als die Frequenz ν_1 des einfallenden Lichts (Abb. 4.1.3-5). Im Falle der Stokesstreuung geht das Atom in einen energiereicheren Zustand über. Das emittierte Photon hat die Frequenz $\nu_2 = \nu_1 - \frac{E_2 - E_1}{h}$. Der umgekehrte Fall die Antistokesstreuung tritt nur ein, wenn sich vor der Einstrahlung ein Teil der Atome im angeregten Zustand E_2 befindet. Die Frequenz der gestreuten Photonen ist $\nu_2 = \nu_1 + \frac{E_2 - E_1}{h}$. Der sogenannte Schwingungs-Ramaneffekt hat eine breite Anwendung in der Molekülspektroskopie gefunden (↑ 5.3.4.). Dabei sind E_1 und E_2 Schwingungszustände der Moleküle.

Die Streuung von Photonen an akustischen Phononen wird als Brillouinstreuung bezeichnet (s. a. 8.1.2.2.). Wie bei der Raman- werden auch bei der Brillouinstreuung in der Streustrahlung Stokes- und Antistokeskomponenten beobachtet. Die Frequenzdifferenz zwischen den Frequenzen der Streustrahlungen und der Erregerstrahlung beträgt $|\Delta\nu| = \nu v \sin(\Theta 2)/c$, dabei sind ν die Frequenz der Erregerstrahlung, v die Schall-, c die Lichtgeschwindigkeit und Θ der Streuwinkel.

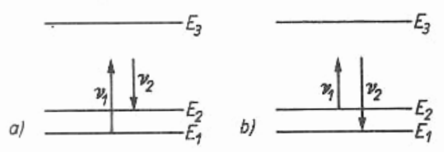


Abb. 4.1.3-5 Ramanstreuung; a) Stokesstreuung, b) Antistokesstreuung

4.1.4. Laser

Laser und Maser sind englische Kunstwörter [*light* bzw. *microwave amplification by stimulated emission of radiation*, engl., Licht- bzw. Mikrowellenverstärkung durch stimulierte Emission von Strahlung]. Laseroszillatoren dienen der Erzeugung von intensiver amplitudenstabiler Strahlung, Laserverstärker sind Anordnungen zur rauscharmen, kohärenten Lichtverstärkung. Die mit Lasern erzeugte und verstärkte Strahlung gleicht im statistischen Verhalten der Strahlung von amplitudenstabilen Radiowellensendern und unterscheidet sich grundlegend von der Wärmestrahlung, die infolge des spontanen Charakters der Emission einer großen Zahl von Atomen durch eine chaotische Verteilung der Feldstärkeamplituden gekennzeichnet ist. Die mit Lasern und Masern aufgebauten Schaltungen und Meßmethoden faßt man in dem Gebiet der *Quantenelektronik* zusammen. Sie ermöglicht es, Schaltungen zur Erzeugung, Verstärkung und Wandlung kohärenter elektromagnetischer Signale bis zum optischen Frequenzbereich, gegenwärtig bis etwa 10^{16} Hz, aufzubauen, während die entsprechenden Methoden der klassischen Elektronik nur bis zu Frequenzen von etwa 10^{10} Hz anwendbar sind. Die Quantenelektronik fußt auf den grundlegenden Arbeiten von Einstein zur *spontanen* und *stimulierten Emission* (↑ 4.1.3.1.) und wurde besonders durch Basov, Prochorov und Townes entwickelt, die für diese Beiträge 1964 mit dem Nobelpreis ausgezeichnet wurden. Die Quantenelektronik – besonders die Laserphysik – hat zu zahlreichen neuen Erkenntnissen über die Natur des Lichts und seine Wechselwirkung mit atomaren Systemen geführt und findet breite Anwendung in Naturwissenschaft und Technik.

4.1.4.1. Lichtverstärkung. Licht kann ähnlich wie ein elektronisches Signal in einer Anordnung mit aktiven Bauelementen verstärkt werden. Eine Lichtwelle der Frequenz ν_0 wird beim Durchlauf durch ein Material, das einen mit

der Energie der Lichtquanten ungefähr übereinstimmenden Energieübergang $E_2 - E_1 = h\nu_{21} \approx h\nu_0$ besitzt (Abb. 4.1.4-1), durch Absorption geschwächt und durch stimulierte Emission verstärkt (↑ 4.1.3.1.). Eine Gesamtverstärkung G = Ausgangsleistung/Eingangsleistung > 1 tritt nur dann auf, wenn sich – abweichend vom thermischen Gleichgewicht – im höheren Energieniveau 2 des aktiven Materials mehr Teilchen als im unteren Niveau 1 befinden;

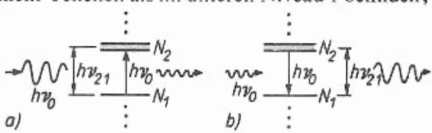


Abb. 4.1.4-1 a) Lichtabsorption bzw. b) Lichtverstärkung beim Durchgang einer Welle durch eine Probe mit $N_2 < N_1$ bzw. $N_2 > N_1$

dann überwiegt die stimulierte Emission die Absorption. Man nennt einen solchen Zustand des Materials *invertiert* und bezeichnet die Differenz der Besetzungszahldichten der (als nicht entartet vorausgesetzten) Niveaus $N_2 - N_1 = \Delta$ als *Besetzungszahlinversionsdichte*. Es gilt $G(\nu) = \exp\left[\Delta B_{12} \frac{h\nu_{21}}{c} g(\nu - \nu_{21}) l\right]$. Danach hängt die Verstärkung von der Besetzungszahlinversion pro Volumeneinheit, dem Einsteinkoeffizienten für stimulierte Emission B_{12}, der *Linienformfunktion* $g(\nu - \nu_{21})$ (↑ 4.1.3.2.) und der Länge l des verstärkenden Materials ab. Bei der Verstärkung durch stimulierte Emission bleibt der Kohärenzzustand der eingestrahlten Welle erhalten. Nur durch die neben der stimulierten Emission stets auftretende spontane Emission entsteht eine zusätzliche Strahlung mit chaotischer Amplitudenverteilung, ein dem verstärkten Signal überlagertes *Rauschen*.

Als *Inversions-* oder *Pumpmechanismen* bezeichnet man Prozesse, mit denen man invertierte Zustände des aktiven Materials erzeugt. Durch Pumpen zwischen den beiden Niveaus 1 und 2, d. h. in einem Zweiniveausystem (↑ Abb. 4.1.4-1), läßt sich stationär keine positive Besetzungszahlinversion erreichen, man erzielt durch starke Einstrahlung von Pumplicht oder durch Teilchenstoß nur einen Zustand mit $N_2 = N_1$, also $\Delta = 0$ und spricht in diesem Fall von einer Sättigung des Übergangs. Dagegen kann man durch starkes Pumpen in einem Drei- oder Vierniveausystem positive Besetzungszahlinversionen herstellen (Abb. 4.1.4-2). Dazu wird das aktive Material durch den starken Pumpprozeß P aus dem Grundniveau in das Hilfsniveau 3 angeregt. Tab. 4.1.4-3 gibt eine Übersicht über einige Pumpprozesse, die zur Anregung der Laser häufig angewendet werden (s. a. 4.3.3.). Aus diesem Niveau gelangt das atomare System durch strahlungslose Prozesse, sogenannte Relaxationsprozesse (↑ 4.1.3.2.2.), sehr schnell in das obere Laserniveau 2. Durch diesen schnellen Übergang von 3 nach 2 bleibt das Niveau 3 nahezu unbevölkert, wodurch der Übergang vom Grundniveau in das Niveau 3 auch bei starkem Pumpen nicht gesättigt werden kann. Infolgedessen können im Grenzfall alle Teilchen in das obere Laserniveau gepumpt werden.

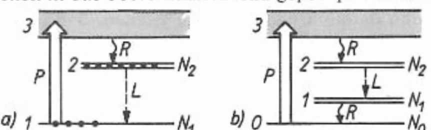


Abb. 4.1.4-2 Herstellung einer positiven Besetzungsinversion, a) in einem Dreiniveau-, b) in einem Vierniveausystem; P, R, L Pump-, Relaxations- bzw. Laserübergänge

Tab. 4.1.4-3 Pumpprozesse

Pumpprozeß	Beispiele für geeignete aktive Materialien
Lichteinstrahlung	Rubinkristall, Neodymglas, Farbstoffe in Lösung
Elektronen-, Atom- und Ionenstöße in einer Gasentladung	Helium-Neon-Gemisch Argon CO_2-Stickstoff-Gemisch
Elektronenstrahl	CO_2, Xenon, GaAs-Halbleiter-Kristalle
Injektion von Ladungsträgern in einen p-n Übergang im Halbleiter	GaAs
chemische Reaktionen	$H_2 + F_2 \rightarrow 2\,HF$

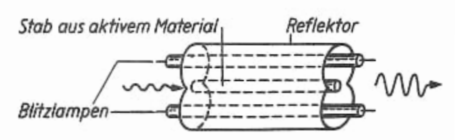


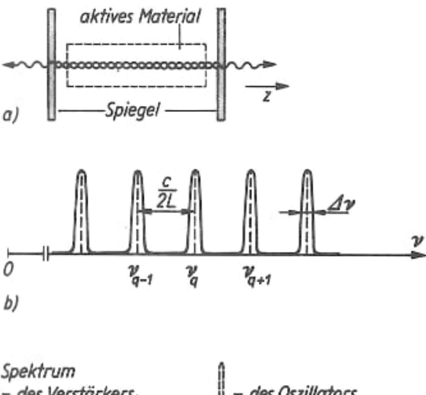
Abb. 4.1.4-4 Optisch gepumpter Laserverstärker

Das Vierniveauschema ist gegenüber dem Dreiniveauschema für die Lichtverstärkung von Vorteil, weil das untere Laserniveau durch strahlungslose Prozesse schnell entvölkert wird; eine positive Besetzungszahlinversion ist deshalb schon bei kleinen Werten von N_2 erreichbar, während im Dreiniveauschema N_2 größer als die Hälfte der Gesamtzahl der aktiven Teilchen sein

muß. Typische Vertreter der Dreiniveau- bzw. der Vierniveaulaser mit optischem Pumpprozeß sind der Rubin- bzw. Neodymlaser. Die Lichtanregung kann durch leistungsstarke Gasentladungsröhren erfolgen, wobei das Pumplicht durch Reflektoren in das aktive Material konzentriert wird (Abb. 4.1.4-4).

4.1.4.2. Laserresonatoren. Zum Aufbau von Lasern, Oszillatoren und allgemein von quantenelektronischen Schaltungen werden außer den Verstärkern Resonatoren benötigt. Der einfachste Resonator für optische Wellen besteht aus zwei parallel justierten, ebenen Spiegeln; man bezeichnet ihn als ebenen *Fabry-Perot-Resonator* (Abb. 4.1.4-5a). Eine sich in der Achsenrichtung z des Resonators ausbreitende ebene Welle wird an den Spiegeln mit dem Reflexionsvermögen R_a bzw. R_b wiederholt reflektiert, wobei bei jeder Reflexion ein bestimmter Teil der Wellen hindurchgelassen wird. Dabei ergeben sich für Licht der Wellenlänge λ_q mit $q \frac{\lambda_q}{2} = l$ (q natürlich) stehende Wellen mit stabilen Feldverteilungen im Resonator, die als *Eigenschwingungen* oder *Moden* des Resonators bezeichnet werden (Abb. 4.1.4-5b, s. a. 2.3.3., 4.1.2.1.2.).
Durch die Auskopplung von Strahlungsenergie an den Spiegeln sowie durch andere Verluste im Resonator wird eine einmal angeregte Eigenschwingung im Resonator gedämpft. Die *Lebensdauer* τ der elektromagnetischen Anregung im Resonator mit hoch reflektierenden Spiegeln ist durch $\tau \approx \frac{l}{c} \frac{1}{1-R}$ mit $R = \sqrt{R_a R_b}$ gegeben. Die Dämpfung führt zu einer Linienbreite $\Delta\nu$ der Eigenschwingungen. Als Güte des Resonators wird die Größe $Q = 2\pi\tau\nu_q$ bezeichnet. Wenn die Wellen im Resonator seitlich nicht unendlich ausgedehnt sind, so ergeben sich kompliziertere Feldverteilungen im Resonator. Die Moden müssen dann außer durch den Index q, der als longitudinaler Modenindex bezeichnet wird, durch zwei transversale Modenindizes charakterisiert werden.
Optische Resonatoren können außer mit ebenen Spiegeln auch mit gekrümmten Spiegeln aufgebaut werden, wobei die Feldverteilung und die Güte der Moden von der Krümmung der Spiegel abhängen. Die aus einem optischen Resonator ausgekoppelte Strahlung einer Mode zeichnet sich auch im freien Raum durch eine hohe räumliche Kohärenz aus, d. h. durch feste Phasenbeziehungen des Felds über den Querschnitt eines Lichtbündels. Insbesondere besitzt ein solches Lichtbündel eine beugungsbegrenzte Divergenz der Größenordnung Wellenlänge/Bündeldurchmesser.

4.1.4.3. Oszillationsbedingung. Setzt man einen Laserverstärker in einen optischen Resonator (↑ Abb. 4.1.4-5a), kann man die Energieverluste der Moden des passiven Resonators teilweise oder vollständig kompensieren. Der Resonator wird entdämpft, und der Verstärker erhält bei den Modenfrequenzen eine Rückkopplung. Durch die Entdämpfung werden die Lebensdauer der Energie im Resonator vergrößert, die Linienbreite der Schwingung verringert und die Güte erhöht. Der Laserverstärker beginnt bei einer bestimmten Eigenfrequenz des Resonators zu oszillieren, wenn das Produkt aus der Verstärkung für einen Hin- und Rücklauf im aktiven Material G und der Rückkopplung durch Reflexion R bei dieser Frequenz den Wert 1 erreicht. Diese Bedingung $G(\nu)R(\nu) = 1$ wird als *Schwellenbedingung* des Lasers bezeichnet. Aus ihr folgen die Besetzungszahlinversion und damit auch die Stärke des Pumpprozesses, die zur Erreichung von Oszillationen erforderlich sind.



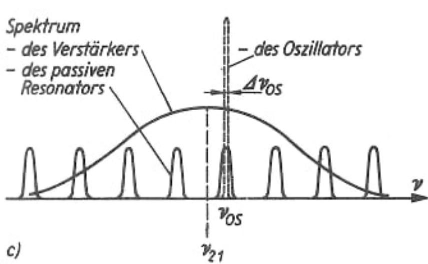


Abb. 4.1.4-5 Fabry-Perot-Resonator; a) Anordnung der Spiegel, b) Modenspektrum in Abhängigkeit von der Frequenz, c) Spektren

4.1.4.4. Linienbreite und Kohärenz. Infolge der fortlaufenden Verstärkung und Rückkopplung entsteht eine Welle hoher Feldstärkeamplitude. Die Feldstärke wird nach oben begrenzt durch den infolge stimulierter Emission mit der Feldstärke stark anwachsenden Abbau der Besetzungszahlinversion und damit der Verstärkung. Diese Nichtlinearität im Verstärkungsprozeß bewirkt die Herausbildung einer Welle mit sta-

biler Amplitude. Die Welle ist durch hohe Monochromasie, d. h. eine sehr geringe Frequenzbreite gekennzeichnet (Abb. 4.1.4-5c). Die kleinste erreichbare Frequenzbreite des Laseroszillators $\Delta\nu_{os}$ wird ebenso wie bei einem elektronischen Generator durch Rauschvorgänge verursacht, die im Falle des Lichts durch die spontane Emission gegeben sind. Die kleinsten bisher erreichten relativen Linienbreiten liegen bei $\Delta\nu_{os}/\nu_{os} \approx 10^{-14}$. Der Laser ist ein rückgekoppelter Verstärker für die spontane Emission und strahlt eine nahezu monochromatische Welle mit *stabiler Amplitude* $\hat{E}$ und langsam schwankender Phase $\varphi(t)$ der Feldstärke aus: $E = \hat{E}\sin[2\pi\nu_{os}t + \varphi(t)]$. Dagegen zeigt die Strahlung einer thermischen Lichtquelle, die man mit geeigneten Filtern auf die gleiche Frequenzbreite $\Delta\nu_{os}$ einengt, eine völlig andere Verteilung der Amplituden; sie können jeden Wert zwischen Null und Unendlich mit einer bestimmten Wahrscheinlichkeit annehmen, wobei die kleinen Werte am häufigsten sind.

Bei den klassischen Interferenzexperimenten nach YOUNG oder MICHELSON unterscheidet sich die Laserstrahlung außer in der Lichtstärke nicht von frequenzgefilterter thermischer Strahlung. Dieser Sachverhalt wird dadurch beschrieben, daß beiden Strahlungen bei gleicher Linienbreite auch die gleiche Kohärenz erster Ordnung zukommt. Der große praktische Vorteil der Laser besteht bei diesen Experimenten darin, daß man die Erzielung hoher räumlicher Kohärenz und kleiner Frequenzbreite nicht mit großen Intensitätsverlusten erkaufen muß. Dadurch konnten auch Interferenzexperimente zwischen 2 unabhängigen Oszillatoren durchgeführt werden, was eine Beobachtungsdauer $\ll 1/\Delta\nu_{os}$ erfordert.

Die Intensitätsinterferometrie nach BROWN und TWISS ergibt grundsätzliche Unterschiede zwischen frequenzgefilterter thermischer und Laserstrahlung. Bei der Intensitätsinterferometrie wird die einfallende Strahlung an einem Spiegel geteilt und auf Lichtwegen, die sich um eine Länge L unterscheiden, auf zwei fotoelektrische Empfänger gegeben. Die Empfängersignale werden elektronisch miteinander multipliziert und das Produkt, das proportional $I(t) \cdot I(t - \tau)$ mit $\tau = L/c$ ist, wird angezeigt. Während das Ausgangssignal bei frequenzgefilterter thermischer Strahlung mit wachsendem τ im statistischen Mittel stark abfällt, ist es bei Laserstrahlung unabhängig von τ. Das thermische Licht ist also durch eine Photonenklumpung gekennzeichnet, die man Photonenbunching nennt. Auf der Grundlage der Intensitätsinterferometrie kann ein verallgemeinerter Kohärenzbegriff eingeführt werden, demzufolge das Laserlicht von höherer Ordnung kohärent ist als thermisches Licht gleicher Bandbreite.

4.1.4.5. Impulslaser. Mit Hilfe von Laseroszillatoren lassen sich kurze Lichtimpulse herstellen.

Eine Modulation der Pumpleistung wird auf die Laserleistung übertragen, wenn die Einschwingdauer des aktiven Materials und des Resonators klein sind. Dieses Prinzip wird z. B. zur elektrischen Modulation von Halbleiterlasern bis zum Gigahertzbereich angewendet.

Eine sogenannte Güteschaltung wird benutzt, um im Lasermaterial über einen langen Zeitraum Energie aufzuspeichern und diese in sehr kurzer Zeit auszusenden. Die Speicherwirkung wird erreicht, indem man während des Pumpvorgangs die Rückkopplung im optischen Resonator klein hält, z. B. durch Schrägstellen eines Spiegels, durch elektrooptische Schalter oder durch zusätzliche Absorber. Nach dem Einsetzen der Laseroszillation wird die Güte sehr schnell erhöht, z. B. durch Drehung eines Spiegels, Veränderung des elektrooptischen Schalters oder infolge der Sättigung des zusätzlichen Absorbers durch die wachsende Leistung der Strahlung selbst. Es können Impulse von einigen 10^{-9} s Dauer erzeugt werden.

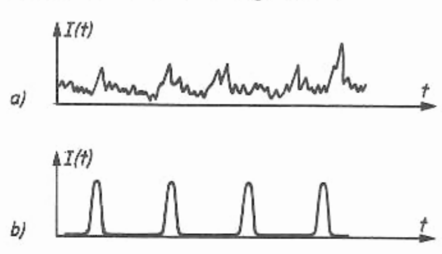


Abb. 4.1.4-6 Abhängigkeit der Laserintensität von der Zeit a) für unsynchronisierte, b) für synchronisierte Moden

Durch *Modensynchronisation* können noch kürzere Impulse hergestellt werden. Dazu wird das Lasermaterial so stark gepumpt, daß der Laser auf vielen Resonatormoden (↑ Abb. 4.1.4-5) oszilliert. Wenn die einzelnen Moden unabhängig oszillieren, sendet der Laser Licht mit einer zeitlich veränderlichen Intensität mit sehr kurzen, unregelmäßigen Intensitätsspitzen aus (Abb. 4.1.4.-6). Die Modensynchronisation kann durch aktive Modulation der Resonatorgüte mit der Modenabstandsfrequenz $c/2L$ oder durch sättigbare Absorber mit schneller Relaxation erfolgen. Es gelingt, periodische Folgen von Impulsen mit einem Abstand $2L/c$ und einer Dauer $2L/pc$ herzustellen, wobei p die Zahl der synchronisierten Moden ist. Es werden Impulse mit Längen im Pikosekundenbereich bei Leistungen im Gigawattbereich erzeugt.

4.1.4.6. Parameter typischer Laser. Die Frequenz des Lasers ist innerhalb der Linienbreite

des aktiven Materials auf die Frequenz des entsprechenden laseraktiven Übergangs festgelegt. Durch die Möglichkeit, aus einer Vielzahl laseraktiver Substanzen auszuwählen, kann man bei nahezu allen Frequenzen zwischen dem ultravioletten und infraroten Spektralgebiet Laser aufbauen. Für eine Frequenzabstimmung eignen sich besonders die Farbstofflaser, weil die entsprechenden Übergänge sehr breit sind. Mit einigen Farbstoffen kann man einen solchen Laser über das gesamte sichtbare Gebiet abstimmen. Die Frequenz von Halbleiterlasern läßt sich durch Druck oder äußere Felder verändern. Außerdem kann mit Methoden der nichtlinearen Optik (↑ 4.1.4.7.4.) eine Frequenzwandlung vorgenommen werden. Leistung, Impulsenergie und Impulsdauer typischer Laser sind in Tab. 4.1.4-7 zusammengestellt (Tafel 4 und 78).

4.1.4.7. Anwendungen von Lasern. Die breiten Anwendungsmöglichkeiten des Lasers resultieren aus der großen Frequenzstabilität, der Monochromasie, der hohen zeitlichen und räumlichen Kohärenz und der hohen Leistung bei beugungsbegrenzter Divergenz.

4.1.4.7.1. *Standards für Länge, Zeit und Frequenz.* Stabilisierte Laser gewährleisten eine höhere relative Genauigkeit der Wellenlänge der Strahlung als der bisherige Längenstandard auf der Basis der Kryptonlampe. Mit einem durch Iodgas stabilisierten He-Ne-Laser kann die Einheit der Länge genauer als 10^{-10} festgelegt werden, statt wie bisher auf 10^{-8}. Die Strahlung des gleichen Lasers kann auch als Frequenz-Standard und Zeit-Standard ausgenutzt werden, wodurch relative Genauigkeiten von 10^{-13} statt bisher 10^{-12} möglich sind. Durch die Messung von Wellenlänge und Frequenz eines Lasers lassen sich Präzisionsbestimmungen der Lichtgeschwindigkeit sowie anderer Naturkonstanten wie der Rydbergkonstante und der Feinstrukturkonstante (↑ 4.2.2.6.) vornehmen.

4.1.4.7.2. *Längenmeßtechnik.* Große terrestrische ($\sim 10^4$ m) sowie kleine kosmische Entfernungen (z. B. Erde – Mond) werden unter Ausnutzung der hohen Impulsleistung und der kleinen Winkeldivergenz der Laser mittels Laserradar vermessen, d. h. durch Messung der Laufzeit kurzer Lichtimpulse. Mittlere terrestrische Entfernungen (10^2 bis 10^4 m) werden durch Phasenmessung einer modulierten Laserwelle mit mm-Genauigkeit bestimmt. Entfernungen zwischen 0 und 10 m mißt man mit hoher Präzision mit Laserinterferometern, wodurch Punkte in drei Dimensionen mit einer Genauigkeit von 0,01 μm festgelegt werden können.

4.1.4.7.3. *Holographie.* Sie ist ein kohärent optisches Abbildungsverfahren, bei dem im Aufzeichnungsmaterial – z. B. einer Fotoplatte – die vom Objekt beeinflußte Laserwelle mit einer Vergleichswelle des gleichen Lasers überlagert wird. Dadurch wird nicht nur die Amplitude, sondern auch die Phase der Objektwelle gespeichert. Mit dem Vergleichsstrahl kann das gespeicherte *Hologramm* ausgelesen werden, wobei ein dreidimensionales Bild des Objekts entsteht. Die Holographie findet u. a. Anwendung bei der optischen Speicherung, der Zeichen-, Objekt- und Fehlererkennung, der optischen Filterung und dem Formvergleich.

Tab. 4.1.4-7 Repräsentative Parameterwerte für einige Lasertypen; λ Betriebswellenlänge, P max. Leistung, E max. Impulsenergie, t Impulsdauer

Lasertyp	Betriebsart	λ in μm	P in W	E in Ws	t in s
He–Ne-L.	kontin.	0,633	10^{-1}	–	–
Ar-Ionen-L.	kontin.	0,514	10	–	–
CO_2-L.	kontin.	10,6	10^5	–	–
	Impuls	10,6	10^{10}	10^4	10^{-6}
Nd–YAG	kontin.	1,06	10^2	–	–
mit Güteschaltung	Impuls	1,06	10^8	10^0	10^{-8}
mit Modensynchron.	Impuls	1,06	10^9	10^{-1}	10^{-10}
Nd-Glas mit Modensynchron.	Impuls	1,06	10^{10}	10^{-1}	10^{-11}
GaAs-L.	kontin.	0,84	10^{-1}	–	–

4.1.4.7.4. *Nichtlineare Optik.* Infolge der hohen Feldstärke der Laserstrahlung, die vergleichbar mit den inneren Feldstärken in den Atomen ist, können in optischen Materialien nichtlineare Prozesse angeregt werden. Durch sie können unter anderem Oberwellen sowie Wellen mit Summen- und Differenzfrequenzen erzeugt werden. Es gelingt z. B., ausgehend von der Strahlung des Neodymlasers bei $\lambda = 1{,}06$ μm Strahlung auf den Wellenlängen 0,53 μm, 0,35 μm, 0,27 μm, 0,21 μm bis herab zu 0,03 μm herzustellen. Beim parametrischen Effekt zerfällt ein Laserphoton mit der Frequenz ν_L und der Wellenzahl k_L in zwei niederfrequente Photonen unter Erfüllung des Energie- und Impulserhaltungssatzes $\nu_L = \nu_1 + \nu_2$ und $k_L = k_1 + k_2$. Auf dieser Basis wird im optischen parametrischen Oszillator Strahlung mit abstimmbarer Wellenlänge erzeugt.

4.1.4.7.5. *Laserspektroskopie.* Mit der hochauflösenden Laserspektroskopie gelingt es unter Ausnutzung der hohen Monochromasie und Frequenzstabilität sehr kleine Frequenzabstände – bis herab zu einigen 100 Hz – zu mes-

sen. Dadurch können Hyperfeinstrukturaufspaltungen und kleinste Frequenzverschiebungen genau bestimmt werden. Dabei ersetzen die *abstimmbaren Laser* nicht nur die Lichtquelle, sondern auch den Monochromator der klassischen Spektrometer. Durch die hohe Empfindlichkeit der Laserspektroskopie können einzelne Atome und Moleküle nachgewiesen und untersucht werden.

Die *Laser-Ultrakurzzeit-Spektroskopie* verwendet leistungsstarke gütegeschaltete und modensynchronisierte Laser zur Anregung und Beobachtung von physikalischen, chemischen und biologischen Elementarprozessen bis herab zum Pikosekundenbereich.

Durch Kopplung spektroskopischer Nachweisverfahren mit dem Laserradar ist eine Spektroskopie über große Entfernungen möglich, was besonders für die Erkennung von Schadstoffquellen im Umweltschutz von Bedeutung ist.

Mit der *Lasermikrospektralanalyse* wird eine hohe räumliche Auflösung – in der Größenordnung einiger Lichtwellenlängen in jeder Dimension – erreicht, was besonders in der Werkstoffprüfung, aber auch in Biologie und Medizin Anwendung findet (Tafel 20/21).

Die selektive spektrale Anregung einzelner Energieniveaus ermöglicht *die Stimulierung von fotochemischen Reaktionen sowie eine Isotopentrennung* (↑ 3.3.1.4.).

4.1.4.7.6. *Informationsübertragung und -verarbeitung.* Mit den Lasern als kohärente Strahlungsquellen extrem hoher Frequenz können große Informationsdichten übertragen werden. In den Schaltungen der integrierten Optik werden Halbleiterlaser mit geeigneten Modulationssystemen gekoppelt und an Lichtfaserleitungen zur Übertragung angeschlossen. Bei langen Übertragungswegen können Laserverstärker eingebaut werden. Mit gekoppelten Lasern lassen sich sehr schnell reagierende logische Elemente und auf dieser Basis eine *optische Informationsverarbeitung* aufbauen.

4.1.4.7.7. *Materialbearbeitung.* Durch die Möglichkeit, hohe Dauer- oder Impulsleistungen von Lasern auf kleine Flächen bzw. in kleine Volumina zu fokussieren, ergeben sich sehr effektive Möglichkeiten zum Schneiden, Bohren und Schweißen. Die bearbeiteten Materialien reichen von Kristallen extremer Härte und Glas über Metalle bis zu Textilien, Kunststoffolien, Papier und biologischem Gewebe (Tafel 79).

Durch die hohe Energiekonzentration der Laserstrahlung können *Plasmen* (↑ 9.1.) sehr hoher Temperatur ($\sim 10^8$ K) erzeugt werden. Daraus folgt prinzipiell die Möglichkeit, mit Lasern die *gesteuerte Kernfusion* (↑ 9.4.) zu erreichen, woran gegenwärtig zur Lösung des Weltenergieproblems in vielen Laboratorien gearbeitet wird.

4.2. Struktur und Energiespektren der Atome

4.2.1. Atome und Quantentheorie

4.2.1.1. Zur Entwicklung des Problems. Das Verständnis der Atomstruktur war in der historischen Entwicklung sehr eng mit der Herausbildung quantentheoretischer Konzeptionen verbunden. Mit dem *Rutherfordschen Atommodell* (↑ 1.1.4.1.) war die Frage nach den im atomaren Bereich geltenden physikalischen Gesetzen gestellt, denn klassische Mechanik und Elektrodynamik können die Stabilität der Atome prinzipiell nicht erklären. Eine nur vorläufige Lösung des Problems brachte die *Bohrsche Theorie* (↑ 1.1.4.4.) mit ihren quantisierten strahlungsfreien Elektronenbahnen und den postulierten Quantensprüngen zwischen ihnen. Mit dieser Theorie konnte man zwar eine Fülle von experimentell beobachteten Spektrallinien theoretisch deuten, sie war aber weder allgemein anwendbar noch logisch geschlossen. Sie benutzte einerseits die klassische Bahnvorstellung und alle Aussagen der klassischen Mechanik für die Bewegungen auf diesen Bahnen, postulierte aber andererseits, daß von allen diesen Bahnen nur einige wenige erlaubt, alle anderen aber verboten sind. Die Auswahl der erlaubten Bahnen geschah durch physikalisch völlig unverständliche Quantenbedingungen (↑ 1.1.4.6.), die sich überdies bei Mehrelektronensystemen gar nicht anwenden ließen; z. B. konnte die Energie des Heliumatoms prinzipiell nicht genau berechnet werden.

Es war daher ein großer Erfolg, als SCHRÖDINGER aus der Theorie der von DE BROGLIE eingeführten Materiewellen (↑ 1.2.3.1.) eine *Wellenmechanik* (↑ 1.2.3.) entwickelte, nach der man die diskreten Energieniveaus der Atome über die Plancksche Beziehung $E = h\nu$ als Eigenfrequenzen stehender Materiewellen physikalisch verstehen konnte. Die endgültige Lösung brachte aber erst die moderne *Quantentheorie*, die unter wesentlicher Benutzung des Wahrscheinlichkeitsbegriffs den Teilchen- und Wellenaspekt verbindet (↑ 2.5.1.3.). Diese Theorie ist logisch konsistent und in sich geschlossen; die anschauliche Vorstellung von Elektronenbahnen ist aber mit ihr nicht mehr vereinbar.

Die Anwendung der Quantentheorie auf den Atombau erklärt im Prinzip alle Erfahrungstatsachen der Atomphysik, z. B. die Existenz und die Energien der diskreten stationären Zustände, die den strahlungsfreien Bohrschen Bahnen entsprechen; die Auswahlregeln für die Emission von Strahlung durch Übergänge zwischen diesen Energieniveaus und die Intensitätsverhältnisse bei mehreren miteinander konkurrierenden Strahlungsübergängen; die Verschiebung und Aufspaltung der Energieniveaus und der Spektrallinien durch die Einwirkung äußerer

elektrischer und magnetischer Felder auf die Atome; das Auftreten von relativ weitreichenden van-der-Waalsschen Kräften zwischen den Atomen, der charakteristischen Kräfte der chemischen Bindung, u. a.

4.2.1.2. Wechselwirkungen im Atom. Ein neutrales Atom besteht aus einem Z-fach positiv geladenen Atomkern und den ihn umgebenden Z einfach negativ geladenen Elektronen; bei Ionen ist die Elektronenzahl von der Kernladungszahl verschieden. Da der Kern sehr klein und sehr viel schwerer als die Elektronen der Hülle ist, kann man ihn in guter Näherung als ruhendes, Z-fach positiv geladenes Kraftzentrum idealisieren. Daraus folgt andererseits, daß es relativ schwierig ist, aus experimentellen Daten der Hüllenphysik mehr über die Atomkerne zu erfahren. Die beherrschenden Wechselwirkungen sind die *elektrostatischen Anziehungskräfte* zwischen den Elektronen und dem Kern sowie die entsprechenden Abstoßungskräfte zwischen den Elektronen. Da das zugehörige Coulombsche Kraftgesetz einfach und genau bekannt ist, beruht die Hüllenphysik in dieser Hinsicht auf einem tragfähigen Fundament. Die wesentlichen Strukturen der Atome und Moleküle und der aus ihnen aufgebauten Substanzen werden erklärbar durch die Kombination quantenmechanischer und elektrostatischer Gesetzmäßigkeiten. Da die Geschwindigkeiten der äußeren Elektronen etwa um einen Faktor 10^2 kleiner als die Lichtgeschwindigkeit sind, können relativistische Effekte meist vernachlässigt werden. Die nichtrelativistische Quantenmechanik, vor allem charakterisiert durch die Schrödingergleichung (↑ 2.5.4.1.), genügt dann zur Behandlung des Atombaus.

Für manche Problemstellungen muß das erläuterte Grundmodell verfeinert werden durch die Berücksichtigung weiterer Wechselwirkungen und Effekte. Die *Mitbewegung des Kerns* hat den größten Einfluß beim Wasserstoff; sie kann dort exakt berücksichtigt werden. Bei schwereren Kernen ist der Mitbewegungseffekt meist unwesentlich. Das *endliche Kernvolumen* und die Wechselwirkung der Elektronen mit elektrischen und magnetischen Momenten des Kerns führen zu den sehr kleinen Effekten der *Hyperfeinstruktur* (↑ 4.2.4.8.). Zur Erklärung der Feinstruktur der Energieniveaus (↑ 4.2.4.4.) muß man über die Näherung von Elektronen mit konstanter Masse und rein elektrostatischer Wechselwirkung zwischen ihnen hinaus *magnetische Kopplungen* und die relativistische Geschwindigkeitsabhängigkeit der Elektronenmasse berücksichtigen. Für das Wasserstoffatom wird diese Aufgabe von der Diracschen Theorie gelöst. Die *relativistischen Korrekturen* sind dabei im allgemeinen klein, werden aber wesentlich für die inneren Elektronen schwerer Atome, da deren Geschwindigkeit mit der Lichtgeschwindigkeit vergleichbar ist. Die Energien der inneren Elektronen sind wichtig für die Röntgenspektren, ihre räumliche Verteilung für die von der Elektronenhülle am Kernort erzeugten elektrischen und magnetischen Felder.

4.2.2. Wasserstoffatom und wasserstoffähnliche Atome

4.2.2.1. Dimensionsanalyse und Abschätzung des Atomradius. Das Elektron im Feld des Wasserstoffkerns befindet sich im Potential $V(r) = -e^2/4\pi\varepsilon_0 r$ der elektrischen Coulombwechselwirkung. Die potentielle Energie hängt nur vom Abstand r zwischen Elektron und Proton ab, sie ist um so kleiner (Minuszeichen!), je näher sich die beiden Teilchen sind, wie es bei Anziehung zwischen ihnen sein muß. Die möglichen Energiewerte und die stationären Elektronenzustände ergeben sich in der Näherung eines ruhenden Kerns aus der Lösung der Schrödingergleichung (↑ 2.5.4.1.) für die Wellenfunktion des Elektrons. Diese nichtrelativistische, aber quantenmechanische Behandlung des Problems enthält folgende physikalische Konstanten: die Plancksche Konstante $\hbar$, die Elektronenmasse m und die Kombination $e^2/4\pi\varepsilon_0$ als Maß für die Stärke der elektrostatischen Anziehung. $\hbar$ hat die Dimension [Energie · Zeit], $e^2/4\pi\varepsilon_0$ die von [Energie · Länge], also der Quotient beider $e^2/4\pi\varepsilon_0\hbar$ die einer Geschwindigkeit. Das dimensionslose Verhältnis dieser atomaren Geschwindigkeitseinheit zur Lichtgeschwindigkeit gibt die sogenannte *Sommerfeldsche Feinstrukturkonstante* (↑ 4.2.2.6.) $e^2/4\pi\varepsilon_0\hbar c \approx 1/137 \approx 10^{-2}$; die Kleinheit dieser Größe rechtfertigt die nichtrelativistische Behandlung des Problems. Da $mv^2/2$ die kinetische Energie eines Teilchens ist, liefert $m(e^2/4\pi\varepsilon_0\hbar)^2 = 27{,}21$ eV eine natürliche Energieeinheit, die als *atomare Energieeinheit* bezeichnet wird. Schließlich erhält man eine Konstantenkombination von der Dimension einer Länge, indem man $e^2/4\pi\varepsilon_0$ durch die Energieeinheit dividiert. Auf diese Weise entsteht der *Bohrsche Radius* $a_0 = \hbar^2/m(e^2/4\pi\varepsilon_0) \approx 0{,}5 \cdot 10^{-10}$ m. Ohne c läßt sich keine dimensionslose Zahl aus den drei Ausgangskonstanten bilden. Daher sind die angegebene Länge, Geschwindigkeit und Energie mit Sicherheit für das Problem charakteristisch, insbesondere bestimmt a_0 die Ausdehnung des Wasserstoffatoms. In der älteren Bohrschen Theorie (s. a. 1.1.4.5.) ergab sich a_0 als kleinster Kreisbahnradius, daraus resultiert die angegebene Bezeichnung. In der modernen Quantentheorie existiert kein scharf definierter Radius, dennoch charakterisiert a_0 die Ausdehnung der Elektronenwolke, genauer die der Wahrscheinlichkeitsverteilung des Elektrons (↑ 4.2.2.4.).

Ohne die Plancksche Konstante, d. h. nur aus m

und $e^2/4\pi\varepsilon_0$, lassen sich überhaupt keine Längen oder Energien bilden. Dies zeigt, daß die klassische Physik die Stabilität der Atome, ihren endlichen Radius, prinzipiell nicht erklären kann.

Ein physikalisches Verständnis für die *energetischen Verhältnisse* vermittelt folgende einfache Abschätzung. Wird das Elektron in einem Bereich vom Radius R lokalisiert, so hat es im Feld des Kerns eine potentielle Energie von der Größenordnung $-e^2/4\pi\varepsilon_0 R$. Andererseits ist mit dieser Lokalisierung entsprechend der Unbestimmtheitsrelation (↑ 2.5.1.7.) das Auftreten von Impulsen der Ordnung $\hbar/R$ verbunden, wobei im Grundzustand diese Größe sicher nicht wesentlich überschritten wird. Daher ist die mittlere kinetische Energie des Elektrons von der Ordnung $m^{-1}(\hbar/R)^2$. Das Minimum der Summe $\hbar^2/mR^2 - e^2/4\pi\varepsilon_0 R$ als Funktion von R liegt bei $R_0 = a_0/2$, wobei der Faktor 1/2 bei dieser groben Abschätzung belanglos ist. Der Atomradius stellt sich also so ein, daß die Summe aus der potentiellen und der als Folge der Unbestimmtheitsrelation auftretenden kinetischen Energie so klein wie möglich wird.

4.2.2.2. Diskussion des Effektivpotentials. Das Coulombpotential ist ein Spezialfall eines allgemeinen kugelsymmetrischen Potentials. Das Problem der Bestimmung der stationären Quantenzustände in solchen Potentialen kann unter Benutzung des Drehimpulssatzes und der quantenmechanischen Eigenschaften des Drehimpulses auf die Lösung eines eindimensionalen Problems für die radiale Koordinate r zurückgeführt werden (↑ 2.5.4.4.). Die radiale Bewegung erfolgt sowohl klassisch wie quantenmechanisch in einem effektiven Potential (s. a. 2.1.6.3., 2.5.4.5.), das sich vom eigentlichen Potential um das zum Drehimpulsquadrat proportionale Zentrifugalpotential unterscheidet. Die Quantisierung des Drehimpulsquadrats $l^2 \to \hbar^2 l(l+1)$, mit $l = 0, 1, 2, \ldots$ führt für das hier interessierende Coulombpotential auf $V_{\text{eff}} = -e^2/4\pi\varepsilon_0 r + \hbar^2 l(l+1)/2mr^2$. Einige qualitative Merkmale der Lösungen der Schrödingergleichung für $l \neq 0$ können bereits aus dem Verlauf dieses drehimpulsabhängigen Effektivpotentials erschlossen werden. Der Dimensionsanalyse des Problems entsprechend fallen alle dimensionsbehafteten Faktoren weg, wenn V_{eff} in atomaren Energieeinheiten, abgekürzt a. E., und r in atomaren Längeneinheiten gemessen wird. Es gilt $V_{\text{eff}}/[me^4/(4\pi\varepsilon_0\hbar)^2] = V_{\text{eff}}/\text{a. E.} = -1/q + l(l+1)/2q^2$ mit $q = r/a_0$. Das wichtigste Merkmal von V_{eff} ist, daß es auch für beliebig große l-Werte noch ein Minimum durchläuft (Abb. 4.2.2-1). Dies folgt aus der großen Reichweite des nur mit q^{-1} abfallenden Coulombpotentials. Die Existenz des Minimums erlaubt die Ausbildung *gebundener Zustände bei beliebig großen Drehimpulswerten*. Diese Besonderheit des Coulombpotentials gilt sowohl für die klassischen Bewegungen wie für die mechanischen Zustände. Das Minimum von V_{eff} liegt bei $q_{\min}(l) = l(l+1)$ und hat die Tiefe $-[2l(l+1)]^{-1}$ a. E., es wird also für große l sehr flach und verschiebt sich weit nach außen. Beides geschieht wegen $l(l+1) \simeq (l+1/2)^2$ für große l-Werte nach einem quadratischen Bildungsgesetz. Da der tiefste Energiewert zum gegebenen Effektivpotential oberhalb von dessen Minimum liegen muß, ist damit eine untere Grenze für die möglichen Energiewerte der gebundenen Zustände mit fixiertem l-Wert gegeben. Der Bereich maximaler Aufenthaltswahrscheinlichkeit des Elektrons muß sich zusammen mit der Lage des Potentialminimums mit wachsendem l zu größeren Kernabständen verschieben (↑ 4.2.2.4.).

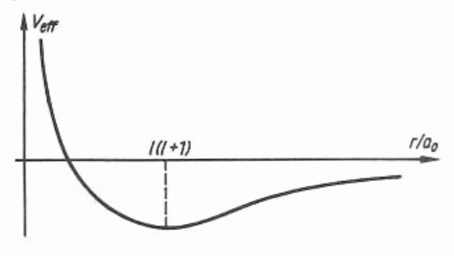


Abb. 4.2.2-1 Verlauf des zum Coulombpotential gehörenden Effektivpotentials für $l \neq 0$

4.2.2.3. Energiespektrum des Wasserstoffatoms. Die möglichen diskreten Energiewerte und die zugehörigen quantenmechanischen Wellenfunktionen können durch die mathematische Lösung der radialen Schrödingergleichung (↑ 2.5.4.5.) bestimmt werden. Die Energien hängen nach der Beziehung

$$(1)\quad E_{n_r l} = -\frac{me^4}{(4\pi\varepsilon_0\hbar)^2}\frac{1/2}{(n_r + l + 1)^2} = -\frac{1/2}{(n_r + l + 1)^2}\ \text{a. E.} \equiv -\frac{1}{2n^2}\ \text{a. E.}$$

von der radialen Quantenzahl $n = 0, 1, 2 \ldots$ (↑ 2.5.4.5.) und der Drehimpulsquantenzahl $l = 0, 1, 2 \ldots$ ab (Abb. 4.2.2-2). Beide durchlaufen unabhängig voneinander alle natürlichen Zahlen. Die Energien der gebundenen Zustände sind sämtlich negativ. Die Energie $E = 0$ hätte ein Elektron, das in großem Abstand vom Kern ruht. $|E_{n_r l}|$ ist daher diejenige Energie, die zur Ablösung des Elektrons, also zur Ionisation des Atoms, mindestens zugeführt werden muß. Je „negativer" die Energie ist, um so fester ist das Elektron an den Kern gebunden. Den niedrigsten Energiewert erhält man für $n_r = l = 0$. Das sind die Quantenzahlen des *Grundzustands* mit der Energie $E_0 = -\frac{1}{2}$ a. E. $\approx -13{,}33$ eV.

Die Bindungs- bzw. *Ionisationsenergie* des Elektrons im Grundzustand des Wasserstoffatoms beträgt +13,33 eV.
Die Energiewerte $E_{n_r l}$ hängen nicht von n_r und l einzeln ab, sondern nur von deren Summe, genauer vom Quadrat der sogenannten *Hauptquantenzahl* $n \equiv n_r + l + 1$. Dem Grundzustand entspricht $n = 1$. Das tiefste Niveau eines Niveaus. Die Drehimpulsquantenzahl l wird in diesem Zusammenhang oft als *Nebenquantenzahl* bezeichnet. In der Atomphysik und Spektroskopie hat es sich eingebürgert, einen Elektronenzustand anstatt durch die sich aus der allgemeinen Theorie ergebenden Quantenzahlen n_r und l durch n und l, also durch Haupt- und Nebenquantenzahl, zu charakterisieren.

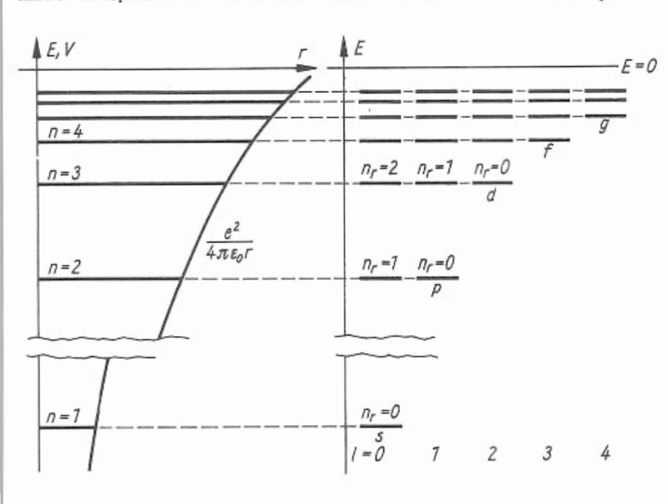


Abb. 4.2.2-2 Lage und Klassifizierung der Energieniveaus eines Teilchens im Coulombpotential

$E_{0l} = -[(2(l+1)^2]^{-1}$ a. E. zu vorgegebenem l-Wert liegt jeweils nur wenig oberhalb von dem Minimum des Effektivpotentials (↑ 4.2.2.2.).
Da der gleiche Wert der Hauptquantenzahl n durch verschiedene Kombinationen von n_r und l entstehen kann, liegt beim Coulombpotential ähnlich wie beim dreidimensionalen isotropen Oszillator (↑ 2.5.4.6.) eine „*zufällige*" *Entartung* von Niveaus mit verschiedenen l-Werten vor. Diese Entartung ist für das Coulombpotential charakteristisch und wird durch jede Abweichung vom Potentialverlauf $V \sim r^{-1}$ aufgehoben. Bei gegebener Energie, also bei gegebener Hauptquantenzahl n, kann l die Werte 0, 1, ..., $n - 1$ annehmen. Zu jedem l-Wert gehören die $(2l + 1)$ verschiedenen Drehimpulseinstellungen, die durch die *Orientierungsquantenzahl* $m = -l, ..., +l$ voneinander unterschieden werden (↑ 2.5.4.4.). Diese $(2l + 1)$-fache Entartung ist allein an die Kugelsymmetrie des Potentials, nicht aber an den speziellen Potentialverlauf gebunden. Durch die hinzukommende zusätzliche Entartung sind nun die Wasserstoffniveaus bei $n = 1$ (Grundzustand) einfach, für $n = 2$ (erster angeregter Zustand) $(1 + 3)$-fach, für $n = 2$ entsprechend $(1 + 3 + 5)$-fach und allgemein n^2-fach entartet. Die Hauptquantenzahl n bestimmt also mit Energiewert und Entartungsgrad die beiden wichtigsten Merkmale

Außerdem werden die l-Werte durch zugeordnete kleine Buchstaben s, p, d, f ... ersetzt (↑ 2.5.4.4.).
Danach kann das Elektron im Wasserstoffatom die Zustände 1*s*; 2*s*, 2*p*; 3*s*, 3*p*, 3*d*; 4*s*, 4*p*, 4*d*, 4*f*; ... annehmen, von denen diejenigen mit gleichem n jeweils gleiche Energie haben. Man verwendet die Hauptquantenzahl n auch zur Kennzeichnung der Niveaus in komplizierteren Atomen (s. a. 4.2.3.2.).

4.2.2.4. Wahrscheinlichkeitsverteilung des Elektrons. Nach den grundlegenden Vorstellungen der Quantenmechanik ist das Elektron nicht strikt lokalisiert, sondern gewissermaßen auf die Umgebung des Kerns verschmiert. Die Dichte der auf diese Weise entstehenden *Elektronenwolke* ist genau als Dichte der räumlichen Aufenthaltswahrscheinlichkeit zu interpretieren (↑ 2.5.1.5.) und als Absolutquadrat der im allgemeinen komplexen Wellenfunktion zu berechnen. Wird also die *Wellenfunktion* $\Phi(\boldsymbol{r})$ als Funktion des Ortes $\boldsymbol{r}$ durch Lösung der Schrödingergleichung bestimmt, so ist damit auch die räumliche Elektronenverteilung bekannt. Wie bei jedem kugelsymmetrischen Potential wird die Wellenfunktion am übersichtlichsten als Funktion sphärischer Polarkoordinaten dargestellt, wobei die Abhängigkeit von den Winkeln ϑ und φ in Gestalt der Kugelfunktionen $Y_{lm}(\vartheta, \varphi)$ bereits vollständig durch die beiden Dreh-

impulsquantenzahlen l und m fixiert wird. ϑ ist der Winkel zwischen Ortsvektor und z-Achse, φ der Drehwinkel um diese Achse. Die Wellenfunktionen des Grundzustands und des ersten angeregten Zustands sind in Tab. 4.2.2-3 angegeben.

Tab. 4.2.2-3 Normierte Wasserstoffwellenfunktionen

Elektronensymbol	Quantenzahlen n	n_r	l	m	Eigenfunktion, $q \equiv r/a_0$
$1s$	1	0	0	0	$[\pi a_0^3]^{-1/2}\, e^{-q}$
$2s$	2	1	0	0	$[8\pi a_0^3]^{-1/2}\,(1 - q/2) \cdot e^{-q/2}$
$2p_z$	2	0	1	0	$[8\pi a_0^3]^{-1/2}\, q/2 \cdot e^{-q/2} \cos\vartheta$
$2p_x$	2	0	1	± 1	$[16\pi a_0^3]^{-1/2}\, q/2 \cdot e^{-q/2} \sin\vartheta\, e^{\pm i\varphi}$

4.2.2.4.1. *Winkelabhängigkeit.* Zu allen s-Zuständen gehören kugelsymmetrische Verteilungen, da die Wellenfunktionen nicht von den Winkeln abhängen. Dagegen sind die Verteilungen zu Zuständen mit von Null verschiedenem Drehimpuls, also beispielsweise von p- oder d-Elektronen, nicht kugelsymmetrisch, bleiben aber wegen des fixierten Werts $\hbar m$ der z-Komponente des Drehimpulses rotationssymmetrisch um die z-Achse (↑ 2.5.3.1.). Als Funktion des Winkels ϑ gegenüber der z-Achse hat die Wellenfunktion und damit auch ihr Absolutquadrat Nullstellen. Die Aufenthaltswahrscheinlichkeit des Elektrons verschwindet dann auf Doppelkegelmänteln mit den entsprechenden Öffnungswinkeln. Für $\vartheta = \pi/2$ entartet der Kegel zur x-y-Ebene. So besteht die Elektronenwolke eines p-Elektrons für $m = 0$ aus zwei durch diese Ebene getrennten Teilen. Solche axialsymmetrischen, *orientierten Verteilungen* führen zum Auftreten bestimmter Valenzwinkel bei der Bildung von Molekülen. Im allgemeinen verschwindet die Wellenfunktion sowohl auf Kugeln als auch auf Kegelmänteln (Abb. 4.2.2-4). Solche Nullflächen werden oft als Knotenflächen bezeichnet.

Für $l \gg 1$ und $|m| \ll l$ zeigen die Verteilungen eine deutliche Orientierung derart, daß das Elektron auf die Umgebung der z-Achse konzentriert ist. Andererseits ist die Wahrscheinlichkeitsdichte für $|m| \approx l \gg 1$ nur in einem kleinen Winkelbereich um $\vartheta = \pi/2$ wesentlich von Null verschieden; die Verteilung konzentriert sich deshalb auf die Umgebung der x-y-Ebene.

Im halbklassischen Bild präzediert für $|m| = l \gg 1$ der Drehimpulsvektor auf einem Kegel mit sehr kleinem Öffnungswinkel (↑ Abb. 2.5.3-1). Dem entspricht eine weitgehende Fixierung der zum Drehimpulsvektor senkrechten

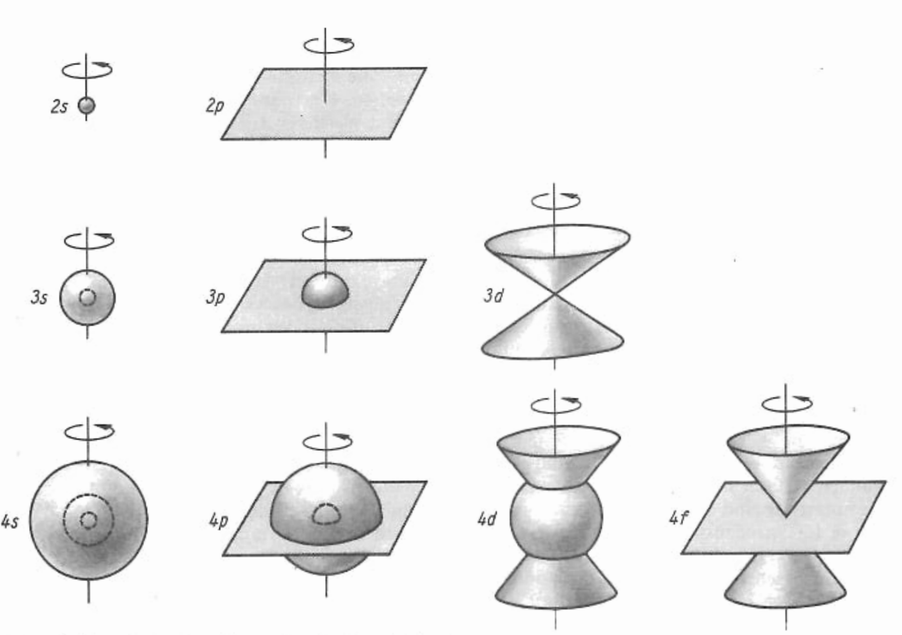


Abb. 4.2.2-4 Knotenflächen für die Zustände mit $m = 0$

klassischen Bahnebene, und dieses Ergebnis korrespondiert mit der erwähnten Konzentration der quantenmechanischen Aufenthaltswahrscheinlichkeit.

4.2.2.4.2. *Radiale Verteilung.* Häufig interessiert man sich für die über alle Winkel gemittelte radiale Wahrscheinlichkeitsverteilung, genauer gesagt für die Wahrscheinlichkeit $w(r)\,dr$, das Elektron in einer dünnen Kugelschale vom Radius r und der Dicke dr anzutreffen. Die Integration der Wahrscheinlichkeitsdichte $|\Phi|^2$ über eine solche Kugelschale betrifft nur das Absolutquadrat $|Y_{lm}(\vartheta, \varphi)|^2$ des von den sphärischen Winkeln abhängenden Anteils der Wellenfunktion. Dieses Integral liefert eine Konstante, so daß mit der Produktdarstellung $\Phi = r^{-1}\,u(r)\,Y_{lm}$ der Wellenfunktion (↑ 2.5.4.5.) $w(r)dr \sim 4\pi r^2 dr \cdot [r^{-1}u(r)]^2$ bzw. $w(r) \sim u^2(r)$ entsteht; $4\pi r^2 dr$ ist das Volumen der Kugelschale. Für s-Zustände mit ihren reellen und winkelunabhängigen Wellenfunktionen ist einfach $w(r) = 4\pi r^2 \Phi^2$. Für das $1s$-Elektron ist danach $w(r) = (4r^2/a_0{}^3)\exp(-2r/a_0)$. Diese Funktion hat ihr Maximum bei $r = a_0$, das ist genau der Radius der Bohrschen Kreisbahn für den Grundzustand. Allgemein zeigt die radiale Verteilung die Ausbildung von Kugelschalen zwischen den radialen Knotenflächen (Abb. 4.2.2-5). Dem $1s$-Zustand entspricht eine diffus auslaufende Kugel. Die Verteilung des $2s$-Elektrons besteht aus einer inneren, bis zur radialen Knotenfläche mit dem Radius $2a_0$ reichenden Innenkugel und einer diese umhüllenden Kugelschale. Nur mit einer Wahrscheinlichkeit von etwa 5% hält sich das Elektron in der inneren Kugel auf. Die höheren s-Verteilungen $3s$, $4s$ usw. bestehen aus mehreren ineinander steckenden, durch Knotenflächen getrennten Kugelschalen und einer Innenkugel. Bei festgehaltenem Drehimpuls (l) werden die radialen Verteilungen mit wachsender Energie (n) breiter und erstrecken sich weiter nach außen. Darin spiegelt sich wider, daß der für die klassische Bewegung erlaubte Bereich von Abständen sich vergrößert. Der minimale Radius verkleinert, der maximale Radius vergrößert sich. Das entgegengesetzte Verhalten zeigen beide Radien, wenn bei festgehaltener Energie (n) der Drehimpuls (l) vergrößert wird. Dementsprechend werden auch die quantenmechanischen radialen Verteilungen schmaler, sie ziehen sich gewissermaßen zusammen. Die radialen Verteilungen zu $n = 3$ und 4 (↑ Abb. 4.2.2-5) zeigen deutlich diese Tendenz. Bei hochangeregten Zuständen mit $n_r = 0$ bzw. $n = l + 1$ wächst $r = n^2 a_0$ quadratisch mit der Hauptquantenzahl. In der älteren Bohrschen Theorie sind dies die Radien der erlaubten Kreisbahnen. Beim Vergleich ist zu beachten, daß die klassischen Kreisbahnen im Zentralfeld gerade diejenigen Bewegungen darstellen, zu denen bei gegebener Energie der größtmögliche Drehimpuls gehört (Tafel 18).

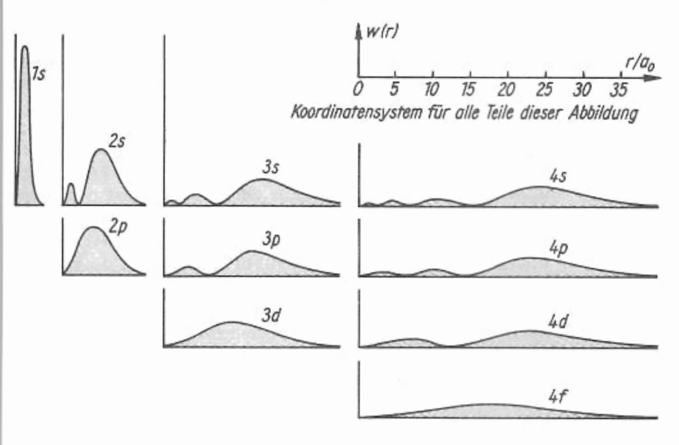


Abb. 4.2.2-5 Radiale Verteilung der Aufenthaltswahrscheinlichkeit des Elektrons im Wasserstoffatom

4.2.2.5. Wasserstoffähnliche Atome und Mitbewegung des Kerns. Wasserstoffähnlich sind alle Systeme, bei denen zwei entgegengesetzt geladene Teilchen durch ihre elektrostatische Anziehung aneinander gebunden sind. Dazu gehören zunächst die nur ein Elektron enthaltenden Ionen der im Periodensystem auf den Wasserstoff folgenden Elemente, also das einfach ionisierte Helium He^+ ($Z = 2$) das zweifach ionisierte Lithium Li^{++} ($Z = 3$) usw. In diesen Fällen besteht der wesentliche Effekt in der Vergrößerung der Kernladung. Der Kern enthält Z Protonen, er hat also die Ladung $+Ze$, Z ist die Kernladungszahl. Bei der Bestimmung der Energieniveaus und Wellenfunktionen des Elektrons ist überall das Quadrat e^2 der Elementarladung durch Ze^2 zu ersetzen. Der Bohrsche Radius $a_0 = 4\pi\varepsilon_0\hbar^2/(me^2)$ geht demnach in a_0/Z über; das bedeutet, daß alle charakteristischen *Längen* der Elektronenwolke und da-

mit diese selbst im stärkeren Anziehungsfeld eines Z-fach positiv geladenen Kerns schrumpfen. Die *Energien* des Wasserstoffspektrums sind proportional zu e^2/a_0 bzw. zu e^4 und daher mit einem Faktor Z^2 zu versehen. Dadurch wird berücksichtigt, daß sich das Elektron im stärkeren Kernfeld (erster Faktor Z) näher am Kern aufhält (zweiter Faktor Z).

Die Kernladung ist zwar die wichtigste, aber nicht die einzige Kerneigenschaft, welche die Elektronenzustände beeinflußt. Die einfachsten weiteren Eigenschaften sind die endliche Masse und das endliche Volumen des Kerns. Die Näherung eines ruhenden Kerns wäre bei unendlich großer Kernmasse exakt. Die *Mitbewegung des Kerns* kann bei echten Einelektronenproblemen (H, He^+, Li^{++}) exakt dadurch berücksichtigt werden, daß man die Elektronenmasse m durch die reduzierte Masse $\mu = mM/(m + M)$ von Elektron und Kern ersetzt (↑ 2.1.6.4.). Dadurch ergeben sich kleine Verschiebungen der Energieniveaus gegenüber den bei Vernachlässigung der Kernbewegung berechneten Werten. Diese Korrekturen sind am größten beim Wasserstoff, jedoch unterscheidet sich auch hier die reduzierte Masse nur um den Faktor $M/(M + m) \approx 1 - m/M \approx 1 - 5 \cdot 10^{-4}$ von der Elektronenmasse. Dennoch sind diese Effekte in den Spektren des emittierten bzw. absorbierten Lichts leicht nachweisbar (↑ 4.3.3.2.2.). Außer beim Wasserstoff überlagert sich dem mit wachsender Kernmasse abnehmenden Mitbewegungseffekt der mit der Größe des Kerns zunehmende *Volumeneffekt* (↑ 4.2.4.9.).

Näherungsweise als wasserstoffähnlich können die inneren Elektronen schwerer Atome behandelt werden. Alle anderen Elektronen der Hülle bilden außerhalb des Systems „Kern-Innenelektronen" eine negativ geladene Kugelschale, die in ihrem Inneren keine elektrostatischen Kräfte hervorruft. Die Innenelektronen verspüren daher genähert nur das Kernfeld.

Wasserstoffartige Bindungszustände sind auch bei Beteiligung anderer Elementarteilchen möglich. So können ein Elektron und ein Positron sich vor der mit der Paarvernichtung verbundenen Zerstrahlung des Systems zu *Positronium* vereinigen; im klassischen Bild umkreisen dabei die beiden gleich schweren Teilchen einander. Die elektrostatische Wechselwirkung ist exakt die gleiche wie beim Wasserstoffatom, jedoch ist die reduzierte Masse nun gleich der halben Elektronenmasse. Dementsprechend ist der Positronium„radius" doppelt so groß wie der des Wasserstoffatoms. Mit dieser Verdopplung der linearen Ausdehnung ist eine Halbierung aller Energien verbunden, die Bindungsenergie des Positroniums beträgt also nur etwa 6,7 eV.

Nachgewiesen ist auch der als *Myonium* bezeichnete Bindungszustand eines Elektrons mit einem positiv geladenen Myon, wobei also das Myon an die Stelle des Protons im Wasserstoffatom tritt. Als einfachstes der sogenannten *Myonenatome*, bei denen im allgemeinen innere Elektronen eines gewöhnlichen Atoms durch negative Myonen ersetzt werden, entsteht das gebundene System Proton – Myon. Hier sind beide Teilchen wesentlich schwerer als das Elektron. Wegen des Massenunterschieds zwischen Elektron und Myon ($m_\mu \approx 207\, m_{el}$) ist dieses System etwa 200mal fester gebunden als das Wasserstoffatom, d. h., sein Radius ist 200mal kleiner und seine Bindungsenergie 200mal größer.

4.2.2.6. Feinstruktur und Spin-Bahn-Kopplung. Aus der genauen Untersuchung des von den Atomen emittierten Lichts (↑ 4.3.3.3.) folgt, daß die Energieniveaus der Atome meist noch eine feine Aufspaltung in mehrere, beim Wasserstoff in zwei dicht beieinanderliegende Niveaus zeigen. Diese Feinstruktur kann nicht durch elektrostatische Wechselwirkungen erklärt werden. Sie beruht auf dem mit dem Elektronenspin (↑ 1.1.2.5.) verknüpften magnetischen Moment. Wie noch erläutert wird, sind Parallel- und Antiparallelstellung von Bahn- und Spindrehimpuls des Elektrons dadurch nicht mehr energetisch gleichwertig.

Das Elektron bewegt sich im elektrischen Feld des Atomkerns. Ein mit dem Elektron mitbewegter Beobachter mißt dann nach den Gesetzen der Elektrodynamik und speziellen Relativitätstheorie auch ein geschwindigkeitsproportionales magnetisches Feld $\boldsymbol{B}$, das bei Umkehr der Bewegungsrichtung sein Vorzeichen ändert und sich insgesamt als zum Bahndrehimpuls $\boldsymbol{l}$ proportional herausstellt. Dieses Feld wirkt auf das mit dem Elektron verbundene, zum Spindrehimpuls $\boldsymbol{s}$ proportionale magnetische Eigenmoment $\boldsymbol{\mu}_s$. Die zugehörige *magnetische Kopplungsenergie* $-\boldsymbol{\mu}_s\boldsymbol{B}$ hat für Parallel- und Antiparallelstellung der beiden Drehimpulse verschiedenes Vorzeichen. Quantitativ erhält man für das radiale elektrische Kernfeld $\boldsymbol{E} = (e/4\pi\varepsilon_0 r^2)\, \boldsymbol{r}/r$, für das Magnetfeld $\boldsymbol{B} = -\boldsymbol{v} \times \boldsymbol{E}/c^2 = el/4\pi\varepsilon_0 mc^2r^3)$, für das magnetische Spinmoment $\boldsymbol{\mu}_s = -(e/m)\, \boldsymbol{s}$ und für

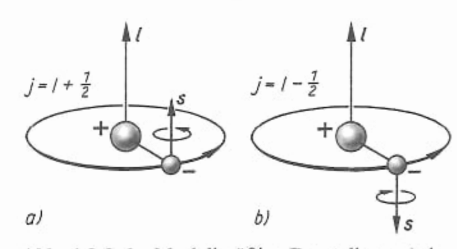


Abb. 4.2.2-6 Modellmäßige Darstellung a) der Parallelstellung, b) der Antiparallelstellung von Bahn- und Spindrehimpuls

die gesamte Spin-Bahn-Kopplungsenergie $E_{SB} = -\mu_s B = [e^2/(4\pi\varepsilon_0 m^2 c^2 r^3)]\, ls$.

Bei Parallelstellung von l und s ergibt sich eine Energieerhöhung, bei Antiparallelstellung eine Erniedrigung (Abb. 4.2.2-6). Bei einem Zustand ohne Bahndrehimpuls verschwindet der gesamte Effekt; s-Zustände werden also nicht aufgespalten. Weitere relativistische Effekte komplizieren das Bild und führen zu einer zusätzlichen kleinen Verschiebung aller, auch der s-Niveaus.

Die Beschränkung auf die beiden durch Parallel- und Antiparallelstellung anschaulich charakterisierten Möglichkeiten folgt aus den grundlegenden quantenmechanischen Aussagen über Drehimpulse (↑ 2.5.3.). Da das Quadrat $\hbar^2 s(s+1)$ des Spindrehimpulses mit $s = 1/2$ festliegt, kann das Quadrat des Gesamtdrehimpulses $j = l + s$ für $l \neq 0$ nur die beiden Werte $\hbar^2 j(j+1)$ mit $j = l - 1/2$ (antiparallel) oder $j = l + 1/2$ (parallel) annehmen (s. a. 2.5.3.4.). Für $l = 0$ stimmt der resultierende Gesamtdrehimpuls des Elektrons mit seinem Spindrehimpuls überein. Die Elektronenzustände können durch die Quantenzahlen n, l und j charakterisiert werden. Es ist üblich, den j-Wert unten rechts an das bisherige Elektronensymbol anzufügen. So bezeichnet $1s_{1/2}$ den Grundzustand, $2s_{1/2}$, $2p_{1/2}$ und $2p_{3/2}$ sind die drei zur Hauptquantenzahl $n = 2$ gehörenden Zustände, bei den p-Zuständen nach $j = l \mp 1/2$ unterschieden.

Die Größenordnung der Energieaufspaltung erhält man durch Einsetzen von $\hbar^2$ für ls und von a_0^3 für r^3:

$$E_{SB} \sim \frac{e^2}{4\pi\varepsilon_0 a_0} \frac{\hbar^2}{m^2 c^2 a_0^2} = \frac{e^2}{4\pi\varepsilon_0 a_0} \left(\frac{e^2}{4\pi\varepsilon_0 \hbar c}\right)^2.$$

Dabei ist der erste Faktor die atomare Energieeinheit. Der zweite Faktor stellt das Quadrat der dimensionslosen *Feinstrukturkonstante* dar (↑ 4.2.2.1.). Da diese genähert den Wert 1/137 hat, sind die Aufspaltungen beim Wasserstoff etwa um den Faktor 10^{-4} kleiner als die Energien der Wasserstoffniveaus, d. h. von der Größenordnung 0,01 eV. Das Auftreten der Lichtgeschwindigkeit bzw. der Feinstrukturkonstante zeigt deutlich, daß der Rahmen einer nichtrelativistischen Beschreibung bei der Erklärung der Feinstruktur verlassen werden muß. Die Kleinheit der Aufspaltungen rechtfertigt andererseits nochmals die vorherige nichtrelativistische Behandlung.

Bei *wasserstoffähnlichen Elektronenzuständen* (↑ 4.2.2.5.) ist auch die Größe der Feinstrukturaufspaltungen mit dem Faktor Z^2 zu multiplizieren. Das ist vor allem für die inneren Elektronen mittlerer und schwerer Atome wesentlich, deren Energieniveaus die nach Herausschlagen solcher Elektronen aus der Atomhülle emittierten Röntgenspektren bestimmen (↑ 1.1.4.7. und 4.3.3.2.3.). Da Z^2 bei schweren Atomen in die Größenordnung von 10^3 bis 10^4 kommt, entstehen nennenswerte Feinstrukturaufspaltungen von einigen 10 eV, die *relativistischen Dubletts*.

Aus der genaueren, von DIRAC entwickelten relativistischen Quantentheorie des Elektrons mit Spin folgt

$$E_{nj} = -\frac{1}{2} \frac{me^4}{(4\pi\varepsilon_0\hbar)^2} \times \left[\frac{1}{n^2} + \left(\frac{e^2}{4\pi\varepsilon_0\hbar c}\right)^2 \left(\frac{1}{j+1/2} - \frac{3}{4n}\right)\right]$$

bei Vernachlässigung höherer Potenzen der Feinstrukturkonstante. Nach dieser Feinstrukturformel werden die Energien aller Zustände etwas gesenkt, und zwar bei festem n die mit dem kleinsten j-Wert, das ist $j = 1/2$, am stärksten. Zustände mit gleichem n und j haben danach auch gleiche Energie, unabhängig von der Quantenzahl $l = j \pm 1/2$ des Bahndrehimpulses.

Dies gilt speziell auch für die beiden Zustände $2s_{1/2}$ und $2p_{1/2}$ mit $n = 2$ und $j = 1/2$. Präzisionsmessungen, die 1950 von LAMB und RETHERFORD mit den Mitteln der Mikrowellentechnik durchgeführt wurden, zeigten auch hier noch eine kleine Aufspaltung, den ***Lambshift***. Dieser kann von der Diracschen Theorie nicht erklärt werden. In Übereinstimmung mit dem experimentellen Ergebnis wird er von der Quantenelektrodynamik vorhergesagt, wobei nun auch das elektromagnetische Feld im Inneren des Atoms quantentheoretisch beschrieben wird.

4.2.3. Atome mit mehreren Elektronen

4.2.3.1. Schalenmodell. Die Quantentheorie ist im Prinzip anwendbar auf Atome mit beliebig vielen Elektronen. Zur Bestimmung der stationären Zustände müßte als Lösung einer Schrödingergleichung eine Wellenfunktion berechnet werden, die in komplizierter Weise von den Koordinaten aller Elektronen abhängt. Die für leichte Atome berechneten numerischen Ergebnisse weichen von den experimentellen Resultaten fast nur durch die Vernachlässigung relativistischer Effekte ab.

Grundlegend für die gesamte Theorie des Atombaus ist das *Modell der unabhängigen Elektronen* (↑ 2.5.5.1.), in dem jedes Elektron seine eigene Wellenfunktion φ und damit seine eigene Wahrscheinlichkeitsverteilung hat. Dabei wird das durch die elektrische Abstoßung bewirkte gegenseitige Ausweichen der Elektronen vernachlässigt, während es sich in der exakten Wellenfunktion widerspiegeln würde. Ähnlich wie beim Wasserstoffatom wird φ aus einer

Schrödingergleichung bestimmt. Die potentielle Energie des Elektrons ist aber nicht von vornherein bekannt, sie enthält vielmehr außer dem Potentialfeld des Atomkerns die mittlere Wechselwirkungsenergie des betrachteten Elektrons mit allen anderen, deren negative Raumladungswolken das Kernfeld teilweise abschirmen.

Das zunächst unbekannte mittlere Potential ist durch die Gesetze der Elektrostatik mit der Form der gesamten „Elektronenwolke" des Atoms verknüpft. Diese Wolke entsteht durch additive Überlagerung der den einzelnen Elektronen zugeordneten Wolken, die sich als Absolutquadrate aus deren Wellenfunktionen ergeben. Diese aber hängen als Lösungen der Schrödingergleichung von dem unbekannten mittleren Potential ab. Aus diesem „Teufelskreis" kann eine Lösung nur durch Iterationsverfahren als numerische Approximation gewonnen werden. Aus einem plausiblen Ansatz für die Wellenfunktion werden das Potential, daraus neue Wellenfunktionen und danach im Wechsel weitere berechnet, bis die in aufeinanderfolgenden Schritten erhaltenen Ergebnisse genähert übereinstimmen. Diese *Methode des selbstkonsistenten Felds* wurde von Hartree entwickelt und von Fock verbessert. Das Verfahren kann auch auf schwere Atome angewendet werden.

Das eben diskutierte Modell voneinander unabhängiger Hüllenelektronen wird meist als Schalenmodell bezeichnet, weil die mit diesem Modell berechnete Elektronendichte im Atom in Übereinstimmung mit experimentellen Resultaten eine charakteristische *radiale Schalenstruktur* zeigt (Abb. 4.2.3-1). Diese Struktur

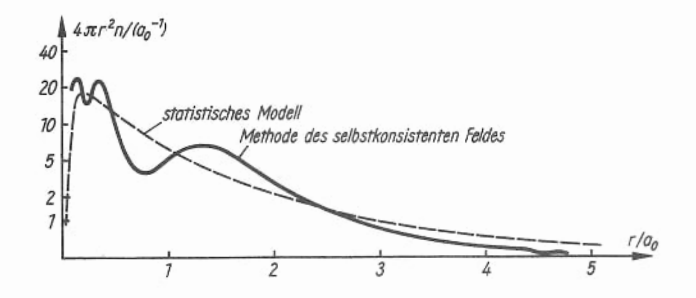


Abb. 4.2.3-1 Radiale Verteilung der Elektronendichte im Argonatom

spiegelt die unterschiedlichen Verteilungen der festgebundenen kernnahen und der weniger fest gebundenen äußeren Elektronen wider. Diese Unterschiede zeigen sich schon beim Vergleich der verschiedenen Elektronenzustände des Wasserstoffatoms (↑ Abb. 4.2.2-5).

4.2.3.2. Elektronenkonfiguration. Auf ein Elektron in unmittelbarer Nähe eines Z-fach geladenen Kerns wirkt das vollständige, nicht abgeschirmte Kernfeld. Auf ein zur Hülle gehörendes Elektron außerhalb des vom Kern und den übrigen $(Z - 1)$ Elektronen gebildeten Atomrumpfs wirkt nur noch das Feld der einen positiven „Überschußladung", alle anderen Ladungen kompensieren sich in ihrer Wirkung nach außen. Stellt man deshalb das Potential in der Form $V(r) = Z(r)\, e^2/(4\pi\varepsilon_0 r)$ dar, so nähert sich die abstandsabhängige Funktion $Z(r)$ für kleine Abstände r der Kernladungszahl Z, für große Abstände der Zahl 1. In einem solchen Potential wird das beim Wasserstoff auftretende zufällige Zusammenfallen von Niveaus mit verschiedenem Drehimpuls l aufgehoben, das Energiespektrum ist vom allgemeinen Typ (↑ 2.5.4.4.).

Dennoch werden die Elektronenzustände im abgeschirmten Kernfeld (s. a. 4.2.3.3.) üblicherweise ebenso wie beim Wasserstoff (↑ 4.2.2.3.) bezeichnet. Die Energie hängt aber nun nicht mehr nur von der Hauptquantenzahl n ab. Am tiefsten liegt stets die Energie des $1s$-Zustands, $2s$ ist der nächste Zustand zu $l = 0$, $2p$ der erste zu $l = 1$ usw.

Das *Pauliprinzip* (↑ 2.5.5.2.) fordert, daß höchstens zwei Elektronen mit entgegengesetzter Spineinstellung die gleichen Quantenzahlen n, l und m haben können. Da die $2l + 1$ verschiedenen, durch $m = -l \dots +l$ charakterisierten Einstellungsmöglichkeiten des Drehimpulses (↑ 2.5.4.4.) im kugelsymmetrischen Kraftfeld alle zum selben Energiewert gehören, können bei Berücksichtigung beider Spineinstellungen $2(2l + 1)$ Elektronen einen Zustand mit definierten Quantenzahlen n und l besetzen, also 2 Elektronen einen s-Zustand, 6 einen p-, 10 einen d-, 14 einen f-Zustand usw. Die beiden Elektronen des Heliumatoms können daher noch im $1s$-Zustand untergebracht werden, während das dritte Elektron im Lithium den nächst höheren, noch unbesetzten $2s$-Zustand einnehmen muß.

Die *Elektronenkonfiguration* eines Atoms gibt an, welche Einelektronenzustände entsprechend dem Pauliprinzip besetzt sind. Sie ist das wichtigste Merkmal der elektronischen Struktur

eines Atoms. Die Elektronenkonfiguration ergibt sich aus der Zahl der Elektronen, die in einem Zustand maximal untergebracht werden können, und aus der energetischen Reihenfolge der Zustände. Die Zahl der in einem Atom wirklich vorhandenen Elektronen mit gegebenem n und l wird durch einen rechten oberen Index am Zustandssymbol angezeigt. Beispielsweise bedeutet die Konfiguration $1s^2 2s^2 2p^2$ des Kohlenstoffatoms, daß mit $n = 1$ zwei s-Elektronen und mit $n = 2$ zwei s- und zwei p-Elektronen vorhanden sind.

Das mittlere Potential des durch die Elektronenwolke abgeschirmten Kernfelds verändert sich von Atom zu Atom nur wenig. Die volle Besetzung eines (nl)-Niveaus und der *Übergang zum energetisch nächsten Niveau* bei Erhöhung der Elektronenzahl um 1 kann sich daher als ein besonderer, von der gleichzeitigen Änderung des Potentialverlaufs nicht kompensierbarer Effekt im Verhalten des Atoms bemerkbar machen. Das wird mit Sicherheit so sein, wenn das neue Niveau sich von dem gerade aufgefüllten energetisch deutlich abhebt. Auf diese Weise führen Abschlüsse bei der Besetzung von Niveaus zu einer charakteristischen *Nichtmonotonie* im Verhalten vieler Atomeigenschaften als Funktion der Elektronenzahl bzw. Kernladungszahl Z, wie sie lange vor der atomtheoretischen Begründung im Periodensystem der Elemente (↑ 4.2.3.4.) formuliert wurde.

4.2.3.3. Elektronen im abgeschirmten Kernfeld. Zur Festlegung der Reihenfolge in der die Niveaus bei schrittweiser Erhöhung der Elektronenzahl besetzt werden, benutzt man oft chemische und spektroskopische Erfahrungen. Physikalisch sind die jeweils hinzukommenden Elektronen „äußere Elektronen" in dem Sinne, daß sie sich bevorzugt in größeren Entfernungen vom Kern als die anderen aufhalten. Die äußeren Elektronen werden bei nicht zu starken Störungen allein angeregt, sie bestimmen die von den Atomen emittierten optischen Spektren und die chemischen Eigenschaften der Substanz.

Am einfachsten ist der Fall der Alkaliatome (Li, Na, K, Rb, Cs) mit nur einem einzigen äußeren *Leuchtelektron*. In einer groben, allerdings unzureichenden Näherung könnte man die Zustände dieses Elektrons denen des Elektrons im Wasserstoffatom gleichsetzen, da der Atomrumpf ebenso wie das Proton beim Wasserstoff einfach positiv geladen ist. Dementsprechend kann die Hauptquantenzahl als erster Anhaltspunkt für die Energie dienen. Tatsächlich ist die Aufenthaltswahrscheinlichkeit für das Gebiet des Atomrumpfs allerdings nicht Null; das äußere Elektron taucht gewissermaßen in den Rumpf ein und unterliegt dort einem stärkeren Bindungspotential, so daß die Energie sinkt bzw. die Bindungsenergie wächst. Dieser Effekt ist um so stärker, je größer der Rumpf, d. h. je schwerer das Atom, und je kleiner der Drehimpuls ist, da Elektronen mit größerem Drehimpuls durch die Zentrifugalkraft vom Kern weggedrückt werden. Ein

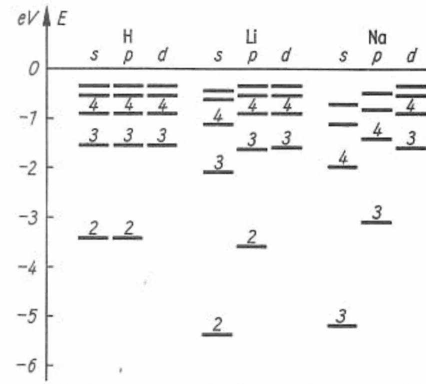


Abb. 4.2.3-2 Energiespektrum des Leuchtelektrons für Wasserstoff, Lithium und Natrium; die Zustände mit $n = 1$ sind nicht eingezeichnet

Vergleich der experimentell aus den optischen Spektren bestimmten Energieniveaus des Leuchtelektrons für Wasserstoff, Lithium und Natrium (Abb. 4.2.3-2) bestätigt das.

Bei kleinen n, genauer bis zu den $3p$-Elektronen, dominieren die großen Energiedifferenzen im Wasserstoffspektrum zwischen Zuständen mit $n = 1, 2$ und 3. Bei den höheren Niveaus werden die durch das Eintauchen der äußeren Elektronen in den Atomrumpf bedingten Energieverschiebungen von gleicher Größenordnung wie die durch Änderung von n bewirkten Energiedifferenzen. Es zeigt sich, daß die Niveaus bei schrittweiser Erhöhung der Kernladungszahl in folgender Reihenfolge besetzt werden: ($1s$), ($2s$, $2p$), ($3s$, $3p$), ($4s$, $3d$, $4p$), ($5s$, $4d$, $5p$), ($6s$, $4f$, $5d$, $6p$), ($7s$, $5f$, $6d$, ...). Dabei erfolgt die Besetzung der d- und f-Niveaus nicht regelmäßig und in Konkurrenz mit den s-Niveaus der gleichen Gruppe. In einer Gruppe sind Zustände zusammengefaßt, die im Bereich der Kernladungszahlen, in dem sie erstmalig besetzt werden, energetisch benachbart sind. Sie bilden eine sogenannte *Energieschale*. Die Niveaus der nächsten Gruppe liegen deutlich höher, so daß Atome mit abgeschlossenen Schalen besonders stabil sind. In der Reihenfolge der angegebenen Gruppen sind das die *Edelgaskonfigurationen* von Helium, Neon, Krypton, Xenon und Radon. Die letzte Gruppe ist nicht vollständig. Die oben angegebenen Gruppen von Zuständen können mit maximal 2, 8, 8, 18, 18 und 32 Elek-

tronen besetzt werden. Die Edelgaskonfigurationen treten daher für $Z = 2, 10, 18, 36, 54$ und 86 auf.
Ein unter Benutzung der gemessenen Ionisationsenergien entworfenes *Niveauschema der Elektronenhülle* (Abb. 4.2.3-3) zeigt nochmals ist besonders stabil, was sich chemisch in der außerordentlichen Reaktionsträgheit des Edelgases Helium und physikalisch in der ungewöhnlich großen *Ionisierungsenergie* 24,56 eV des Heliumatoms ausdrückt. Verglichen mit der Ionisierungsenergie 13,54 eV des Wasserstoffs

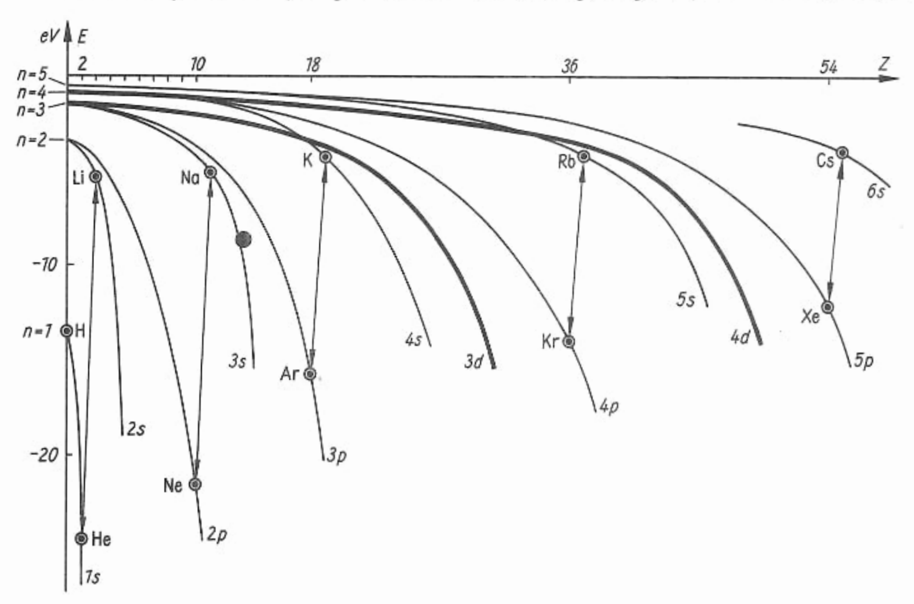


Abb. 4.2.3-3 Niveauschema der Elektronenhülle

den Einfluß der verschieden starken Abschirmung der Kernladung bei *s*-, *p*-, *d*- und *f*-Zuständen und macht die Bildung von Energieschalen deutlich. Damit die Abbildung nicht unübersichtlich wird, ist als letzter nur noch der 6*s*-Zustand eingezeichnet, mit dem die große 32er Schale beginnt. Je näher die Elektronen am Kern, je fester sie mithin gebunden sind, um so mehr verliert der „Taucheffekt" an Bedeutung; die innersten Elektronen zeigen in guter Näherung wieder die beim Wasserstoff vorliegende Entartung von Zuständen mit verschiedenem Drehimpuls (↑ 4.2.2.5.) Insgesamt sinken mit wachsender Kernladungszahl Z alle Energieniveaus zu tieferen Energien ab, und zwar am stärksten die *s*-Zustände (↑ S. 730/731).

4.2.3.4. Atomtheoretische Diskussion des Periodensystems. Das einzige Elektron des Wasserstoffatoms befindet sich im Grundzustand des Atoms im 1*s*-Zustand. Auch alle folgenden Betrachtungen beziehen sich nur auf die Elektronenkonfiguration des Grundzustands. Das beim *Helium* hinzutretende zweite Elektron führt zur Auffüllung des 1*s*-Niveaus, die innerste Schale ist damit komplett. Diese Konfiguration ist beim Helium also fast doppelt so viel Energie notwendig, um eines der beiden Elektronen abzutrennen. Der Grund dafür ist, daß das jeweils andere 1*s*-Elektron die um eine Elementarladung erhöhte Kernladung nur teilweise abschirmt, so daß sich jedes der beiden Heliumelektronen zwar in einem schwächeren Anziehungspotential als das eine Elektron im Heliumion He^+ (↑ 4.2.2.5.), aber in einem stärkeren als das Wasserstoffelektron befindet. Schirmt jedes der beiden Elektronen den Bruchteil ze einer positiven Elementarladung im Kern ab, so ist die *effektive Kernladungszahl* $(2 - z)\,e^2$ an Stelle von $2e^2$. Nach der Wasserstofformel ist dann das Verhältnis der Ionisierungsenergien durch $E_I^{He}/E_I^H = (2 - z)^2/1^2 = 24{,}56\ \text{eV}/13{,}54\ \text{eV} = 1{,}81$ bestimmt, woraus man den plausiblen Wert $z = 0{,}65$ erhält. Das dritte Elektron beim Lithium muß den energetisch nächsten 2*s*-Zustand annehmen. Mit der Erhöhung der Hauptquantenzahl beim Übergang von 1*s* zu 2*s* ist eine deutliche Verschiebung der maximalen Aufenthaltswahrscheinlichkeit nach außen verbunden (↑ Abb. 4.2.2-5). Es ist also das neue Elektron, verglichen mit den

beiden Elektronen der 1*s*-Innenschale, wirklich ein äußeres Elektron. Die auf drei Elementarladungen vergrößerte Kernladung wird durch die beiden Innenelektronen weitgehend abgeschirmt, so daß das Leuchtelektron im Lithium zwar wegen des „Taucheffekts" stärker gebunden ist als das 2*s*-Elektron im Wasserstoff (↑ Abb. 4.2.3-2), aber doch wesentlich schwächer als das 1*s*-Elektron des Wasserstoffs. Die gemessene Ionisierungsenergie von nur 5,37 eV bestätigt die plötzliche Bindungslockerung beim Beginn der neuen Schale. Die effektive Abschirmung des Kernfelds, $3 - z$ antstatt der Kernladungszahl 3, wird durch die erneute Anwendung der Wasserstofformel, analog zur Behandlung des Heliumatoms bestätigt. Aus $E_I^{Li}/E_I^H = (3 - z)^2/2^2/1^2 = 5{,}37\,\text{eV}/13{,}54\,\text{eV} = 0{,}396$ folgt $z = 1{,}74$ also $z/2 = 0{,}87$. Während danach ein 1*s*-Elektron in bezug auf das zweite 1*s*-Elektron der gleichen Schale einen 65%igen Abschirmeffekt hervorruft, wirkt die Ladung jedes dieser Elektronen im Lithium für das 2*s*-Elektron der nächsten Schale schon zu 87% abschirmend. Die schwache Bindung des äußeren Lithiumelektrons äußert sich in einer außerordentlichen Reaktionsfreudigkeit des Alkalimetalls Li bei chemischen Reaktionen. Bei zahlreichen Kombinationen mit anderen Atomen ist mit der Molekülbildung ein Energiegewinn verbunden, weil das im Atom schwach gebundene Elektron im Molekül fester eingebaut werden kann.

Das bei Beryllium hinzutretende vierte Elektron kann ebenfalls noch im 2*s*-Niveau untergebracht werden. Ähnlich wie beim Helium wird die erhöhte Kernladung durch die abschirmende Wirkung des anderen 2*s*-Elektrons nur teilweise kompensiert, so daß auch die Bindung der Außenelektronen verglichen mit Lithium stärker wird. Dementsprechend wächst die Ionisierungsenergie von 5,37 eV für Li auf 9,48 eV für Be. Die festere Bindung drückt sich auch in einem kleineren Atomvolumen des Berylliums aus. Das fünfte Elektron beim *Bor* muß den 2*p*-Zustand als den energetisch nächsten einnehmen. Die damit verbundene Bindungslockerung drückt sich zwar in der verkleinerten Ionisierungsenergie von 8,4 eV aus, ist aber zu schwach, um bei Beryllium von einem Schalenabschluß mit deutlichen Auswirkungen für das qualitative Verhalten der Atome zu sprechen. Der energetische Unterschied zwischen den 2*s*- und 2*p*-Elektronen ist gering. Mit der allmählichen Auffüllung des sechs Elektronen fassenden 2*p*-Niveaus ist eine Verstärkung der Bindung verbunden, die sich in der Zunahme der Ionisierungsenergie von 8,4 eV bei Bor auf 21,5 eV bei *Neon* mit der Konfiguration $1s^22s^22p^6$ ausdrückt, das ist fast so viel wie beim Helium. Neon ist aufgrund seiner besonders stabilen Elektronenkonfiguration wieder ein Edelgas. Die Schale der Elektronen mit der Hauptquantenzahl 2 kann insgesamt 8 Elektronen aufnehmen, dem entspricht die von Lithium bis Neon reichende erste kleine Periode des Periodensystems.

Der Aufbau der dritten Schale ist völlig analog: Mit dem Alkalimetall *Natrium* ($Z = 11$) beginnt die Besetzung des 3*s*-Niveaus, und mit dem Edelgas *Argon* ($Z = 18$) ist das 3*p*-Niveau gefüllt. So entsteht die zweite, wieder aus 8 Elementen bestehende kleine Periode des Periodensystems. Dem gleichartigen elektronischen Aufbauprinzip in beiden Schalen entspricht die weitgehende chemische Ähnlichkeit der einander zugeordneten Elemente.

Beim weiteren Aufbau der Atomhüllen werden nach den 3*p*-Elektronen zunächst die 4*s*-Elektronen eingebaut. Die vierte Schale beginnt bei $Z = 19$ mit dem Alkalimetall Kalium. Die Energie der 4*s*-Elektronen sinkt durch ihr stärkeres „Eintauchen" in den aus den Elektronen der ersten drei Schalen gebildeten positiven Atomrumpf soweit ab, daß sie noch vor den 3*d*-Elektronen eingebaut werden. Diese meiden aufgrund ihres größeren Drehimpulses ($l = 2$) das kernnahe Gebiet, ihre Verteilungen erstrekken sich aber andererseits auch nicht soweit nach außen wie die der 4*s*-Elektronen (↑ Abb. 4.2.2-5). Daher sind die 3*d*-Elektronen, obwohl sie unmittelbar nach den 4*s*-Elektronen eingebaut werden, für die „Oberflächeneigenschaften" des Atoms von geringerer Bedeutung. Das führt z. B. zur chemischen Ähnlichkeit der Elemente Fe, Co und Ni der Eisengruppe, die sich nur durch die Anzahl der 3*d*-Elektronen unterscheiden. Die zehn 3*d*-Elektronen werden von Scandium ($Z = 21$) bis Zink ($Z = 30$) eingebaut. Diese Elemente sind den ersten großen, d. h. aus 18 Elementen bestehenden Perioden des Periodensystems, den sogenannten Nebengruppen, zugeordnet. Würde man sie nicht beachten, so bildeten die verbleibenden Hauptgruppenelemente mit ihren 4*s*- und 4*p*-Elektronen eine Achterperiode, die genau so wie die beiden vorangegangenen aufgebaut wäre. Die vierte Schale endet mit dem *Edelgas Krypton*. Die nächste Schale ist ganz analog aufgebaut: Mit dem Alkalimetall Rubidium beginnt bei $Z = 37$ die Besetzung des 5*s*-Niveaus, von $Z = 39$ bis $Z = 48$ wird das 4*d*-Niveau aufgefüllt, und bei $Z = 54$ wird beim *Edelgas Xenon* das sechste und damit letzte 5*p*-Elektron eingebaut.

In der nächsten, aus den 6*s*-, 4*f*-, 5*d*- und 6*p*-Zuständen bestehenden Schale, treten erstmals *f*-Elektronen auf. Diese sind deutlich weiter innen als die *s*- und *p*-Elektronen der gleichen Schale konzentriert. Das erklärt die fast vollständige Ähnlichkeit der *seltenen Erden*, bei denen die insgesamt vierzehn 4*f*-Zustände aufgefüllt werden. Die insgesamt 32 Elektronen

fassende sechste Schale beginnt bei $Z = 55$ mit dem Alkalimetall Caesium und endet mit dem *Edelgas Radon* bei $Z = 86$. Damit sind die stabilen Elemente fast erschöpft. Die nächste Schale ist unvollständig, nach dem experimentellen Material, das einschließlich der künstlich hergestellten Transurane vorliegt, ist sie analog wie die vorangegangene aufgebaut.

Bei den schweren Elementen können äußere Elektronen oft so leicht angeregt werden, daß sich die Elektronenkonfiguration beim Einbau des Atoms in Moleküle ändern kann. Daher gibt der Elektronenkatalog zwar eine grundsätzliche Orientierung über die chemischen Eigenschaften, diese können jedoch aus ihm nicht eindeutig entnommen werden (↑ S. 730 bis 731).

Der Aufbau einer neuen Schale beginnt stets mit einem s-Elektron, das sehr locker gebunden ist. Auf diese Weise erhält man die *Alkaliatome* Li, Na, K, Rb und Cs mit ihren charakteristischen Einelektronenspektren und ihrer Einwertigkeit in allen chemischen Verbindungen. Die plötzliche Bindungslockerung bei Beginn einer neuen Schale beruht darauf, daß die Verteilung des neuen Elektrons in größerer Entfernung vom Kern konzentriert ist als die der anderen, inneren Elektronen, die die Kernladung für das neue Elektron vollständiger abschirmen, so daß das bindende Kraftfeld relativ schwach ist. Demgegenüber wird beim weiteren Aufbau der neuen Schale jedes weitere Elektron im allgemeinen fester gebunden als das vorhergehende, weil dieses die zusätzliche Ladung im Kern nur teilweise abschirmt. Außer bei den die Schalen abschließenden Edelgasen treten große Ionisierungsenergien, also besonders stabile Konfigurationen, auch bei einigen anderen Atomen auf, bei denen gewissermaßen nur Teilschalen abgeschlossen werden. Das ist sehr deutlich ausgeprägt bei Zink ($Z = 30$), Palladium ($Z = 46$) und Quecksilber ($Z = 80$).

4.2.3.5. Mittlere Abhängigkeit von der Elektronenzahl. Während sich Atomeigenschaften, die durch die relativ wenigen äußeren Elektronen bestimmt werden, dem Periodensystem der Elemente entsprechend nichtmonoton mit der Kernladungszahl Z ändern, zeigen andere Eigenschaften als Funktion von Z einen monotonen Verlauf, z. B. die gesamte, überwiegend von den inneren Elektronen bestimmte Bindungsenergie, die Größe der aus den abgeschlossenen Schalen gebildeten Atomrümpfe, die für die Röntgenspektren (↑ 4.3.3.2.3.) wesentlichen Energieniveaus der inneren Elektronen. Um für Größen wie die Gesamtenergie und die Atomgröße bzw. die Dichteverteilung den über die Schalenstruktur gemittelten Verlauf in Abhängigkeit von Z zu ermitteln, haben THOMAS, FERMI u. a. das *statistische Atommodell* entwickelt. Dabei wird angenommen, daß sich bei hinreichend großer Elektronenzahl bereits in kleinen Raumgebieten mit näherungsweise konstantem Potential viele Elektronen befinden, und daß in jeder dieser Zellen etwa die gleichen Verhältnisse herrschen wie in einem räumlich homogenen Elektronengas mit derjenigen Dichte, die in der gerade betrachteten Zelle vorliegt. Für ein homogenes Elektronengas liefert das Modell der unabhängigen Teilchen den Zusammenhang $p^2/2m|_{\max} = \varkappa\hbar^2 n^{2/3}/2m$ mit $\varkappa = (3\pi^2)^{2/3} \approx 9{,}57$ zwischen der kinetischen Energie der schnellsten Elektronen und der Teilchendichte n. Diese Beziehung ist eine direkte Folge des Pauliprinzips, nach dem sich nicht alle Elektronen in Zuständen mit kleinen Impulsen, d. h. mit geringer kinetischer Energie, befinden können. Im Atom ist nun der Dichteverlauf so zu bestimmen, daß die aus der potentiellen Energie $V(r)$ und der zu $n^{2/3}$ proportionalen kinetischen Energie zusammengesetzte Gesamtenergie der energiereichsten Elektronen für alle Raumpunkte den gleichen Wert hat. Anderenfalls würde Energie gewonnen, wenn bei Änderung der Dichte Elektronen von Orten mit großer Energie zu solchen mit kleinerer Energie übergehen. Das Potential $V(r)$ wird durch die Kernladung und die negative elektrische Raumladungsdichte der Elektronen nach den Grundgesetzen der Elektrostatik erzeugt. Die Durchführung dieses Verfahrens liefert für die Elektronendichte neutraler Atome einen Verlauf, der sich genähert durch die Formel $n(r) = (Z\lambda^{3/2}/16\pi r^{3/2})\, e^{-\sqrt{\lambda r}}$ mit $\lambda \approx 0{,}3/(a_0 Z^{1/3})$ darstellen läßt (↑ Abb. 4.2.3-1). Die Größe λ^{-1} hat etwa die Bedeutung eines Atomradius, der mit $Z^{1/3}$ wächst. Aus dem mittleren Dichteverlauf erhält man die Gesamtenergie der Atomhülle zu $E_{\text{ges}} = -0{,}772 \cdot Z^{7/3}$ a. E. $= -21 \cdot Z^{7/3}$ eV.

Das statistische Modell ist zwar nicht sehr genau, gibt aber den allgemeinen Verlauf befriedigend wieder.

4.2.4. Multiplettstruktur der Energieniveaus

Im Schalenmodell unabhängiger Elektronen wird die elektronische Struktur eines Atoms durch seine Elektronenkonfiguration charakterisiert. Alle Zustände mit gleicher Konfiguration (↑ 4.2.3.2.) haben danach gleiche Energie; eine Anregung des Atoms erfordert eine Veränderung seiner Konfiguration. Dieses Modell beschreibt richtig die niedrig angeregten Zustände solcher Atome, bei denen außerhalb vollständig gefüllter Niveaus nur das einzige Leuchtelektron angeregt wird. Sofern sich aber mehrere Elektronen in teilweise gefüllten Niveaus befinden, kann die Struktur des Energiespektrums nicht allein mit dem Modell unabhängiger Elektronen verstanden werden.

4.2.4.1. Entartung bei teilweise gefüllten Niveaus. Die Kennzeichnung eines Elektronenzustands als $1s$, $2p$ usw., bzw. durch die Hauptquantenzahl n und die Nebenquantenzahl l, (↑ 4.2.2.3.) ist nicht erschöpfend: Es sind nur die den zugehörigen Energiewert bestimmenden Quantenzahlen erfaßt. Die $(2l + 1)$ verschiedenen Einstellmöglichkeiten $m = -l \ldots +l$ des Bahndrehimpulses und die zwei Spineinstellungen $m_s = \pm 1/2$ wurden bisher nur zur Abzählung der in einem gegebenen Energieniveau maximal unterzubringenden Elektronen berücksichtigt.
Ist ein Niveau vollständig besetzt, so ist dadurch eine eindeutige Kennzeichnung des vorliegenden Zustands gegeben: Alle Drehimpuls- und Spineinstellungen werden von je einem Elektron angenommen. Befindet sich nur ein Elektron in einem sonst leeren Niveau, so kann es jeden der insgesamt $2(2l + 1)$ entarteten Zustände einnehmen, das sind sechs Zustände für ein p-Elektron. Für zwei p-Elektronen mit gleicher Hauptquantenzahl, sogenannte *äquivalente p-Elektronen*, ist der Entartungsgrad 15, das ist die Zahl der Möglichkeiten, aus sechs Zuständen zwei als besetzt bzw. drei als unbesetzt auszuwählen (Abb. 4.2.4-1). Entsprechend gibt es für eine p^3-Konfiguration $\binom{6}{3} = 20$, für p^4 wieder 15, für p^5 6 Möglichkeiten, und mit p^6 ist das Niveau gefüllt. Die benutzte Abzählvorschrift trägt der Identität der Elektronen voll Rechnung, denn die äquivalenten Elektronen werden

Konfiguration	$m_s = 1/2$: $m = -1$	0	$+1$	$m_s = -1/2$: $m = -1$	0	$+1$	Entartungsgrad
leer							1
p		↑					6
p^2			↑	↓			15
p^3	↑			↓	↓		20
p^4		↑	↑		↓	↓	15
p^5	↑	↑		↓	↓	↓	6
voll	↑	↑	↑	↓	↓	↓	1

Abb. 4.2.4-1 Besetzungsmöglichkeiten bei äquivalenten p-Elektronen

nicht numeriert, es wird nur zwischen besetzten und unbesetzten Zuständen unterschieden (↑ 2.5.5.3.). Bei vier oder fünf Elektronen ist es offenbar am einfachsten, die Zahl der Möglichkeiten für das Nichtbesetzen zweier Zustände bzw. eines Zustands abzuzählen. Zustände mit k Elektronen in einem teilweise gefüllten Niveau haben den gleichen Entartungsgrad wie solche mit k bis zur vollständigen Besetzung fehlenden Elektronen, sogenannten Löchern.
Nichtäquivalente Elektronen, die sich also in verschiedenen Niveaus befinden, können beim Abzählen der verschiedenen Besetzungsmöglichkeiten als unterscheidbar behandelt werden. Jede Möglichkeit eines Elektrons kann mit jeder der anderen Elektronen kombiniert werden, ohne dadurch das Pauliprinzip zu verletzen. So gibt es bei zwei nichtäquivalenten p-Elektronen, etwa einem $2p$- und einem $3p$-Elektron, insgesamt $6 \cdot 6 = 36$ verschiedene Besetzungsmöglichkeiten, anstatt 15 bei zwei äquivalenten p-Elektronen. Durch das Pauliprinzip und die Ununterscheidbarkeit der Elektronen wird also der Entartungsgrad wesentlich, bei mehr als zwei äquivalenten Elektronen sogar drastisch verringert.

4.2.4.2. Gesamter Bahn- und Spindrehimpuls. Die im bisherigen Modell der Elektronenstruktur auftretenden Entartungen (↑ 4.2.4.1.) sind z. T. modellbedingt. In Wirklichkeit spalten die entarteten Energieterme in mehrere Energiewerte auf. Dafür sind prinzipiell zwei Effekte verantwortlich: die im Modell unabhängiger Elektronen nur grob berücksichtigte *elektrostatische Wechselwirkung zwischen den Elektronen* und die *Spin-Bahn-Kopplung*, die bereits beim Wasserstoff (↑ 4.2.2.6.) zur Feinstruktur führt. Werden beide Kopplungen berücksichtigt, so haben die Bahn- und Spindrehimpulse der einzelnen Elektronen keine konstanten Werte mehr, der Zustand der Elektronenhülle kann dann also auch nicht durch entsprechende Drehimpulsquantenzahlen gekennzeichnet werden. Im anschaulichen Bild tauschen die Elektronen bei aufgrund der Wechselwirkungen zwischen ihnen unvermeidlichen „Stößen" außer Energie und Impuls auch Drehimpuls miteinander aus. Nur der aus allen Bahn- und Spindrehimpulsen bestehende *Gesamtdrehimpuls* der Elektronenhülle bleibt – bei Vernachlässigung der Kopplung mit dem Kernspin (↑ 4.2.4.8.) – eine Erhaltungsgröße, die quantenmechanisch durch ihr Quadrat $\hbar^2 I(I + 1)$ und ihre z-Komponente $\hbar M$, $M = -I \ldots +I$, zu kennzeichnen ist. Die exakten Energieniveaus der Hülle haben nur die $(2I + 1)$-fache Entartung, die den verschiedenen Einstellungen des Gesamtdrehimpulses entspricht.
Im allgemeinen, besonders ausgeprägt bei den leichten Atomen, sind die elektrostatischen Wechselwirkungen wesentlicher als die Spin-Bahn-Kopplungen, d. h., die durch sie bewirkten Niveauaufspaltungen sind deutlich größer als die Energiedifferenzen der Feinstruktur (↑ 4.2.4.4.). Vernachlässigt man dementsprechend zunächst alle Spin-Bahn-Kopplungen, so sind der *gesamte Bahndrehimpuls* $\boldsymbol{L} = \sum_i \boldsymbol{l}_i$ und der *gesamte Spindrehimpuls*

$S = \sum_i s_i$ der Hüllenelektronen für sich allein Erhaltungsgrößen. Völlig analog zu den entsprechenden Drehimpulsen der einzelnen Elektronen oder zum Gesamtdrehimpuls sind dann die Betrags- und Orientierungsquantenzahlen L, M und S, M_s zur Kennzeichnung des Zustands brauchbar. Analog zu den Einelektronenzuständen werden den Werten von $L = 0, 1, 2, 3 \ldots$ die Buchstaben $S, P, D, F \ldots$ zugeordnet (s. a. 2.5.4.4.).

Klassisch erhält man die Vektoren $\boldsymbol{L}$ und $\boldsymbol{S}$ durch vektorielle *Addition der Einzeldrehimpulse*. Quantenmechanisch ergeben sich die bei einer gegebenen Elektronenkonfiguration möglichen L- und S-Werte aus den Regeln für die Addition von Drehimpulsen (↑ 2.5.3.4.). Bei nichtäquivalenten Elektronen können, diese ohne Einschränkungen angewendet, alle Einstellungen von Bahn- und Spindrehimpulsen unabhängig voneinander verwirklicht werden. So ergibt sich bei zwei nichtäquivalenten p-Elektronen, also $s_1 = s_2 = 1/2$ und $l_1 = l_2 = 1$ sofort $S = 0$ und 1; $L = 0$, 1 und 2, also insgesamt die durch (L, S) gekennzeichneten Zustände (0, 0), (1, 0), (2, 0); (0, 1), (1, 1), (2, 1). Es ist üblich, die sogenannte Multiplizität $2S+1$ (↑ 4.2.4.4.) links oben an das Termsymbol zu schreiben. Die obigen Zustände heißen dann also ^1S, ^1F, ^1D; ^3S, ^3P, ^3D.

Bei äquivalenten Elektronen ergeben sich durch Ununterscheidbarkeit und Pauliprinzip wesentliche Einschränkungen. In einem *vollständig* gefüllten Niveau gibt es zu jedem Elektron mit bestimmten Werten von m und m_s ein anderes mit den entgegengesetzten Orientierungen $-m$ und $-m_s$; so daß $M = M_S = 0$ gilt; dazu kann nur ein ^1S-Zustand mit $L = S = 0$ gehören. *Abgeschlossene Schalen* bzw. Teilschalen tragen nicht zu den Gesamtdrehimpulsen bei. Bei einem Elektron oder einem Loch in einem teilweise gefüllten Niveau werden daher L und S allein durch dieses bestimmt, d. h. $L = l$, $S = s = 1/2$; z. B. befindet sich das Atom bei einem oder bei fünf p-Elektronen im Zustand $L = 1$, $S = 1/2$, also in einem ^2P-Zustand.

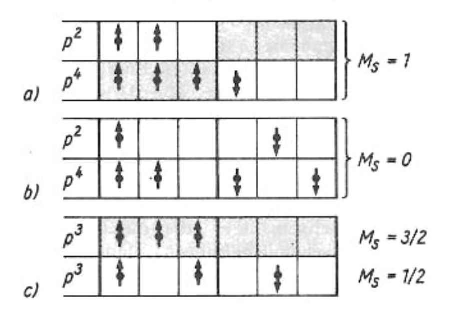


Abb. 4.2.4-2 Spinaddition äquivalenter p-Elektronen; volle oder leere Halbschalen *gerastert*

Sollen sich bei zwei äquivalenten p-Elektronen bzw. bei vier, also bei zwei Löchern, die Spinkomponenten zu $M_s = \pm 1$ addieren, so muß jeweils eine „Halbschale" leer bzw. voll sein (Abb. 4.2.4-2a). Für M sind dann nur die Werte $-1, 0, +1$ entsprechend dem Loch bzw. Elektron in der teilweise gefüllten „Halbschale" möglich. Daher kann $S = 1$ nur mit $L = 1$ kombiniert werden: Es gibt einen ^3P-Zustand. Kompensieren sich dagegen die Spinkomponenten zu $M_s = 0$ (Abb. 4.2.4-2b), so unterscheiden sich die beiden Elektronen bzw. Löcher durch ihre Spinkomponente. Die Addition ihrer Bahndrehimpulse unterliegt keinerlei Einschränkungen, und es treten alle Zustände zu $L = 0, 1, 2$ auf. Die zu $L = 1$ komplettieren den ^3P-Zustand, zu dem oben nur die Zustände mit $M_s = \pm 1$ auftraten. Der Rest liefert wegen $M_s = 0$ nur Zustände mit $S = 0$,

Tab. 4.2.4-3 Mögliche Zustände bei äquivalenten s- und p-Elektronen

Konfiguration	s	p, p^5	p^2, p^4	p^3
Zustände	^2S	^2P	$^1S, {}^1D, {}^3P$	$^2P, {}^2D, {}^4S$

also einen ^1S- und einen ^1D-Zustand. – Bei drei p-Elektronen können die maximalen Werte $M_s = \pm 3/2$ (Abb. 4.2.4-2c) nur durch volle Besetzung der jeweiligen „Halbschale" realisiert werden. Es ist $M = 0$, also auch $L = 0$; der große Spinwert erlaubt keine verschiedenen Kombinationsmöglichkeiten der beiden Bahndrehimpulse. Bei $M_s = \pm 1/2$ befindet sich in der einen Halbschale ein Elektron und in der anderen ein Loch, deren Bahndrehimpulse frei kombinierbar sind. Die Übersicht (Tab. 4.2.4-3) zeigt, daß die bei Konfigurationen von p-Elektronen auftretenden Zustände eindeutig durch ihre (L, S)-Werte charakterisiert werden können. Bei größeren Bahndrehimpulsen und mehr als zwei äquivalenten Elektronen können L-S-Kombinationen auch mehrfach auftreten.

Der *Gesamtspin* ist stets ganzzahlig bei geraden und halbzahlig bei ungerader Elektronenzahl. Sein Maximalwert ist gleich der halben Anzahl aller Elektronen, bzw. aller Löcher bei mehr als halb gefüllten Niveaus, entsprechend der vollständigen Parallelstellung der Einzelspins aller Elektronen bzw. Löcher.

4.2.4.3. Aufspaltung durch Coulombwechselwirkung. Die Zustände mit verschiedenen L-S-Werten besitzen verschiedene Energien, auch wenn sie zur gleichen Elektronenkonfiguration gehören. Solange die Energiedifferenzen klein bleiben gegenüber den Anregungsenergien, die

mit einer Änderung der Konfigurationen verbunden sind, stellt das Modell unabhängiger Elektronen eine brauchbare Ausgangsnäherung dar. Das heißt genauer, daß die aus den entsprechenden Einteilchenwellenfunktionen aufgebauten Wellenfunktionen der gesamten Elektronenhülle genähert anstelle der unbekannten exakten Wellenfunktionen benutzt werden können. Schon die grundlegende Forderung nach Antisymmetrie der Gesamtwellenfunktion (↑ 2.5.5.4., 2.5.5.5.) zwingt dazu, in Form von Slaterdeterminanten *Linearkombinationen einfacher Produktwellenfunktionen* zu benutzen. Am Beispiel des *Ausweichens von Elektronen* mit parallelem Spin (↑ 2.5.5.6.) wurde diskutiert, daß dadurch die gegenseitigen Aufenthaltswahrscheinlichkeiten der Elektronen verändert werden. Daher hängt auch die mittlere elektrostatische Wechselwirkungsenergie der Elektronen von den Spineinstellungen ab. Da die Antisymmetrisierung der Wellenfunktion sich auf den Austausch von Elektronen im Argument der Wellenfunktion bezieht, werden die damit verbundenen Energieunterschiede oft als Austauschenergie bezeichnet. Bereits dieser Effekt allein würde zu einer energetischen Aufspaltung einer Konfiguration führen. Die Notwendigkeit, Wellenfunktionen zu vorgegebenen L-S-Werten aufzubauen, zwingt im allgemeinen zur Kombination mehrerer Slaterdeterminanten. Dadurch werden die mittleren Wechselwirkungsenergien weiter verändert.

Relativ einfach ist die Situation bei den unteren Anregungszuständen von Helium (Abb. 4.2.4-4). Der Grundzustand hat die Konfiguration $(1s)^2$, ist also ein ^1S-Zustand. Die unteren angeregten Zustände entstehen, wenn eines der beiden Elektronen im $1s$-Zustand verbleibt, während das andere in höhere Zustände übergeht. Für solche Konfigurationen können sich die Spins der beiden nichtäquivalenten Elektronen entweder parallel ($S = 1$) oder antiparallel ($S = 0$) einstellen. Im ersten Fall weichen sich die Elektronen gegenseitig aus (↑ 2.5.5.6.). Dies ist wegen der Coulombschen Abstoßung zwischen beiden Elektronen energetisch vorteilhaft. Bei $S = 0$ haben die beiden Elektronen mit antiparallelem Spin die Tendenz, sich benachbart aufzuhalten, wodurch die Energie zunimmt. Bei gleicher Konfiguration liegt die Energie zu $S = 1$ tiefer als die zu $S = 0$.

Die Heliumzustände mit $S = 0$ kombinieren in den Spektren nicht mit denen zu $S = 1$. Das Spektrum besteht daher aus zwei unabhängigen Termsystemen. Vor Aufstellung der zugehörigen Theorie glaubte man zunächst, daß es sich um zwei verschiedene Heliumarten handle, die man *Ortho- und Parahelium* nannte: Sie entsprechen den Zuständen mit $S = 1$ und $S = 0$. Der Grundzustand gehört in dieser Bezeichnungsweise zum Parahelium. Ähnliche Verhältnisse findet man bei den Erdalkalien, bei denen sich ebenfalls zwei Elektronen im dann durch die inneren Elektronen abgeschirmten Kernfeld bewegen.

Auch bei komplizierteren Atomen mit mehr als zwei Elektronen außerhalb voll besetzter Niveaus ist der Zustand mit dem maximalen Spinwert S energetisch bevorzugt, weil in diesem

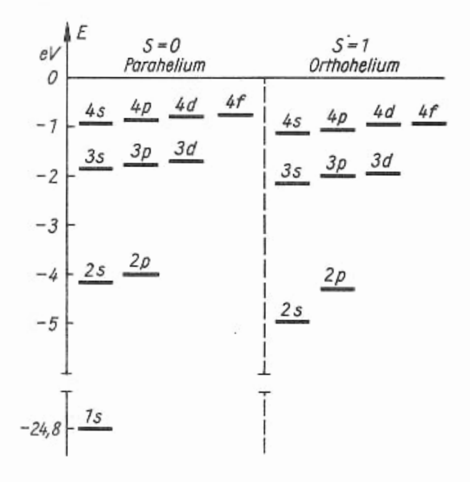


Abb. 4.2.4-4 Energieniveaus des Heliumatoms; die Zustandssymbole $2s$, $2p$ usw. geben den Zustand des angeregten Elektrons an

Zustand die Anzahl der Elektronen mit parallelem Spin am größten ist. Da sich Elektronen mit parallelem Spin gegenseitig ausweichen, wird dadurch die potentielle Energie erniedrigt. Damit wird die *erste Hundsche Regel* verständlich, die durch Analyse des empirischen Materials gefunden wurde. Sie lautet: Bei gegebener Elektronenkonfiguration hat der Grundzustand den größtmöglichen Wert von S.

Falls es mehrere Zustände mit diesem maximalen S-Wert gibt, gilt die *zweite Hundsche Regel*: Zum Grundzustand gehört der größte Wert von L, der mit dem maximalen S-Wert verträglich ist.

4.2.4.4. Feinstruktur. Mit dem Übergang vom einheitlichen Energiewert einer bestimmten Elektronenkonfiguration zu den Energien der einzelnen L-S-Zustände ist der Einfluß der elektrostatischen Wechselwirkungen zwischen Elektron und Kern sowie der Elektronen untereinander vollständig erfaßt. Ähnlich wie beim Wasserstoff (↑ 4.2.2.6.) führen auch bei komplizierteren Atomen magnetische Wechselwirkungen und relativistische Korrekturen zu einer Spin-Bahn-Kopplung. Diese äußert sich in kleinen Energieunterschieden zwischen Zuständen, bei denen Bahn- und Spindrehimpuls

verschieden zueinander orientiert sind. An die Stelle von l und s für das eine Elektron beim Wasserstoff treten die entsprechenden Gesamtdrehimpulse $\boldsymbol{L}$ und $\boldsymbol{S}$ mit den Betragsquantenzahlen L und S. Alle Zustände, die sich bei festen Werten von L und S nur durch die relative Orientierung von $\boldsymbol{L}$ und $\boldsymbol{S}$ zueinander unterscheiden, bilden ein *Multiplett*. Die kleinen Energieunterschiede zwischen den einzelnen Komponenten eines Multipletts bezeichnet man als Feinstruktur. Die Anzahl der zu einem Multiplett gehörenden Zustände verschiedener Energie heißt *Multiplizität*.

Die vektorielle Addition von $\boldsymbol{L}$ und $\boldsymbol{S}$ ergibt den Gesamtdrehimpuls $\boldsymbol{I}$ der Elektronenhülle. Die möglichen Werte von $\boldsymbol{I}^2$ sind $\hbar^2 I(I+1)$; dabei ist I mindestens gleich dem Betrag $|L-S|$ der Differenz bei Antiparallelstellung und höchstens gleich der Summe $L+S$ bei Parallelstellung von $\boldsymbol{L}$ und $\boldsymbol{S}$ (↑ 2.5.3.4.). Für $S=1/2$ gibt es nur diese beiden Fälle, sonst aber noch eine Reihe von Zwischenorientierungen, wobei für benachbarte Orientierungen $\Delta I=1$ ist. Für $S \leqq L$ sind das insgesamt $(2S+1)$ Möglichkeiten, für $S \geqq L$ sind es $(2L+1)$. Diese Zahlen geben also die Multiplizität an. Üblicherweise bezeichnet man $(2S+1)$ als Multiplizität, obwohl diese Zahl nur für $S \leqq L$ diese Bedeutung hat. Das vollständige *Termsymbol* enthält links oben die Multiplizität $(2S+1)$ und rechts unten den Wert von I, z. B. $^4D_{5/2}$ für $L=2$, $S=3/2$, $I=5/2$. I wird häufig als innere Quantenzahl bezeichnet. Zu ungerader Elektronenzahl gehören gerade Multiplizitäten $2S+1$, zu gerader Elektronenzahl ungerade. Diese Folge des halbzahligen Elektronenspins ist als *spektroskopischer Wechselsatz* bekannt.

Die *magnetische Wechselwirkungsenergie* ist analog zu den Verhältnissen beim Wasserstoffatom proportional zum Skalarprodukt $\boldsymbol{LS}$, für das man $2\boldsymbol{LS}/\hbar^2 = I(I+1) - L(L+1) - S(S+1)$ erhält, weil sich aus $\boldsymbol{I}=\boldsymbol{L}+\boldsymbol{S}$ durch Quadrieren $\boldsymbol{I}^2=\boldsymbol{L}^2+\boldsymbol{S}^2+2\boldsymbol{LS}$ ergibt und die Drehimpulsquadrate durch ihre vorgegebenen Werte $\hbar^2 I(I+1)$ usw. ersetzt werden können. Da L und S für alle Terme eines Multipletts konstant sind, gibt $E_{mg}=AI(I+1)$ allein die Aufspaltung eines durch L und S charakterisierten Energieniveaus an. Für benachbarte Terme eines Multipletts ist $\Delta I=1$ oder $E_I - E_{I-1} = 2AI$. Die Energiedifferenz benachbarter Niveaus ist also proportional zum Gesamtdrehimpuls I. Das ist die *Landésche Intervallregel*. Die Größe A nimmt mit steigender Kernladungszahl Z stark zu, so daß die Multiplettaufspaltungen bei schweren Atomen größer werden. A kann positiv oder negativ sein. Im ersten Fall ist die Antiparallel-, im zweiten die Parallelstellung von $\boldsymbol{L}$ und $\boldsymbol{S}$ energetisch bevorzugt; man spricht von normalen und umgekehrten Multipletts. Bei normalen Multipletts liegt der Zustand mit $I=|L-S|$ am tiefsten, bei umgekehrten der mit $I=L+S$. Für die Grundzustände der Atome gilt: Falls es in der Konfiguration nur ein teilweise gefülltes Niveau gibt, so ist $A \geqq 0$ bei höchstens halbgefülltem, $A<0$ bei mehr als halbgefülltem Niveau.

Am Beispiel zweier äquivalenter p-Elektronen (Abb. 4.2.4-5) ist die Mannigfaltigkeit der aus

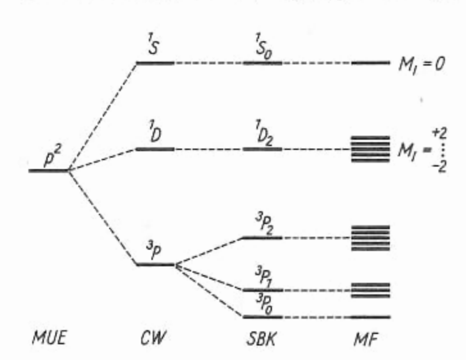


Abb. 4.2.4-5 Multiplettstruktur und Zeemaneffekt für zwei äquivalente p-Elektronen; *MUE* Modell unabhängiger Elektronen, *CW* Coulombwechselwirkung, *SBK* Spin-Bahn-Kopplung, *MF* äußeres Magnetfeld

einer einzigen Konfiguration des Modells unabhängiger Elektronen hervorgehenden Energieniveaus schematisiert. Die Reihenfolge der Niveaus ^1S, ^1D und ^3P entspricht den beiden Hundschen Kopplungsregeln. Durch die Feinstrukturaufspaltung entsteht aus dem ^3P-Zustand ein normales Multiplett mit $I=0$, 1, 2.

Das Kopplungsschema, bei dem zunächst alle Einzeldrehimpulse zu $\boldsymbol{L}$ und $\boldsymbol{S}$ und dann erst diese zu $\boldsymbol{I}$ zusammengesetzt werden, heißt *L-S-* oder *Russell-Saunders-Kopplung*. Es entspricht der bei Atomen meist zutreffenden Annahme, daß die genauere Berücksichtigung der elektrostatischen Wechselwirkungen, die zur Aufspaltung in Niveaus mit verschiedenen Werten von L und S führt, wesentlicher ist als die Beachtung der Kopplungen, die die Feinstrukturaufspaltungen jedes einzelnen L-S-Niveaus liefern. Diese Aufspaltung muß danach klein sein gegen die Energieunterschiede zwischen Niveaus mit verschiedenen L-S-Werten. Diese Voraussetzung ist vor allem bei den leichten Atomen sehr gut erfüllt. Im entgegengesetzten Grenzfall werden zuerst die magnetischen Kopplungen zwischen Bahn- und Spindrehimpuls jedes einzelnen Elektrons berücksichtigt, d. h., es werden wie beim Wasserstoff $\boldsymbol{l}_i$

und s_i zum Gesamtdrehimpuls j_i jedes einzelnen Elektrons addiert und erst dann werden die einzelnen j-Vektoren zum Gesamtdrehimpuls I des Atoms kombiniert. Diese j-j-Kopplung tritt bei Atomen in reiner Form nicht auf.

4.2.4.5. Magnetisches Moment der Atome. Ein System ohne resultierenden Drehimpuls ist völlig kugelsymmetrisch, in ihm gibt es keine ausgezeichnete Richtung. Verschwindet daher der Gesamtdrehimpuls der Elektronenhülle, $I = 0$, so hat das Atom auch kein permanentes von den Elektronen hervorgerufenes magnetisches Moment. Solche Atome sind diamagnetisch. Die magnetischen Momente der Atomkerne (↑ 3.3.2.4.) sind um drei Größenordnungen kleiner, so daß die von ihnen hervorgerufenen Effekte meist vernachlässigt werden können. Elektronenhüllen mit $I \neq 0$ haben ein permanentes magnetisches Moment, sie sind paramagnetisch. Betrachtet man das Atom, genauer seine Elektronenhülle, zunächst als ein hinsichtlich seiner inneren Struktur nicht näher analysiertes System, so ist der Vektor $\boldsymbol{\mu}$ seines magnetischen Moments proportional zum Gesamtdrehimpuls I (↑ 2.2.2.5.): $\boldsymbol{\mu} = gI(-e/2m)$. Dabei ist $-e$ die Ladung und m die Masse eines Elektrons, g der gyromagnetische oder kurz g-Faktor. Die Quantelung des Drehimpulses überträgt sich auf das magnetische Moment, insbesondere sind $\mu_z = -(ge\hbar/2m)\,M$ mit $M = -I \ldots +I$ die $2I + 1$ verschiedenen Werte, die eine hier wieder als μ_z bezeichnete Komponente von $\boldsymbol{\mu}$ annehmen kann.

Der g-Faktor hat weder für alle Atome noch für alle Quantenzustände eines Atoms den gleichen Wert, obwohl immer die Hüllenelektronen das magnetische Moment hervorrufen. Diese zunächst überraschende Tatsache beruht auf dem unterschiedlichen g-Faktor für Bahn- und Spindrehimpulse. Für das mit dem Bahndrehimpuls eines Elektrons bzw. irgendeines

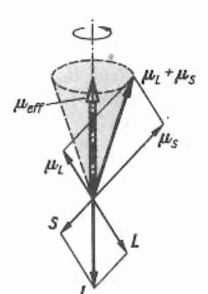


Abb. 4.2.4-6 Halbklassisches Bild zum magnetischen Momen der Elektronenhülle eines Atoms

anderen geladenen Teilchens verbundene Moment ist $g = 1$ (↑ 2.2.2.5.), für das magnetische Spinmoment des Elektrons ist in guter Näherung $g = 2$. Der g-Faktor eines Atoms muß also $g = 1$ sein, wenn der Gesamtdrehimpuls allein durch die Elektronenspins hervorgerufen wird, also für $I = L$, $S = 0$; dagegen ist $g = 2$ für $I = S$, $L = 0$. Im allgemeinen wird g zwischen 1 und 2 liegen. Der richtige Wert von g kann bereits aus der halbklassischen Vorstellung abgeleitet werden, daß die Vektoren $\boldsymbol{L}$ und $\boldsymbol{S}$ von Bahn- und Spindrehimpuls und mit ihnen die Vektoren der ihnen zugeordneten magnetischen Momente $\boldsymbol{\mu}_L$ und $\boldsymbol{\mu}_S$ um die Richtung des resultierenden Gesamtdrehimpulses $\boldsymbol{I}$ präzedieren (Abb. 4.2.4-6). Dabei sind alle zur Drehachse senkrechten Komponenten von $\boldsymbol{L}, \boldsymbol{S}, \boldsymbol{\mu}_L$ und $\boldsymbol{\mu}_S$ im Mittel Null, und nur die Komponenten in Richtung von $\boldsymbol{I}$ bleiben bestehen. Der in der Abbildung mit $\boldsymbol{\mu}_{\text{eff}}$ bezeichnete Vektor hat die Richtung des Gesamtdrehimpulses und entspricht damit dem oben einfach als $\boldsymbol{\mu}$ bezeichneten Vektor. Die Länge dieses Vektors ist gleich der Summe der Achsenprojektionen von $\boldsymbol{\mu}_L$ und $\boldsymbol{\mu}_S$. Als Ergebnis entsteht die Formel $g = [3I(I+1) + S(S+1) - L(L+1)]/[2I(I+1)]$ für den sogenannten *Landéschen g-Faktor*. Die oben diskutierten Grenzfälle $g = 1$ und $g = 2$ sind leicht erkennbar.

4.2.4.6. Zeemaneffekt. Befindet sich das Atom in einem homogenen äußeren Magnetfeld $\boldsymbol{B}$, so kann seine magnetische Wechselwirkungsenergie $-\boldsymbol{\mu B} = -\mu_z B = +(ge\hbar/2m)\,BM_I$ entsprechend den $2I + 1$ verschiedenen Drehimpulsorientierungen, $M_I = -I \ldots +I$, die angegebenen diskreten Werte annehmen. Die bei den bisherigen Fragestellungen willkürliche z-Richtung ist nun durch die Magnetfeldrichtung physikalisch bestimmt, die verschiedenen Einstellungen des Drehimpulses und des mit ihm verbundenen magnetischen Moments beziehen sich auf diese Richtung. Durch die Wechselwirkungsenergie im Magnetfeld spaltet jeder einzelne Term eines Feinstruktur-Multipletts mit $I \neq 0$ nochmals in $2I + 1$ äquidistante Niveaus auf (↑ Abb. 4.2.4-5). Wegen $|\Delta M_I| = 1$ ist der energetische Abstand benachbarter Niveaus $\Delta E = ge\hbar B/2m$; er ist proportional zum angelegten Feld und zum g-Faktor. Die Aufspaltung der Energieniveaus im Magnetfeld führt zu entsprechenden Aufspaltungen der Spektrallinien des beim Übergang zwischen zwei durch ihre I-, L- und S-Werte charakterisierten Energieniveaus emittierten bzw. absorbierten Lichts. Diese Erscheinung wird nach ihrem Entdecker als Zeemanneffekt bezeichnet. Das Aufspaltungsbild ist im allgemeinen relativ kompliziert, da die miteinander kombinierenden Niveaus verschiedene g-Faktoren, also verschieden starke Aufspaltungen im gleichen Magnetfeld, haben. Einfach ist die Situation dann, wenn beide Niveaus gleiches $g_1 = g_2$ haben, das ist insbesondere für $S_1 = S_2 = 0$ und für $L_1 = L_2 = 0$ der Fall. Im ersten Fall der Kombination zweier Singulett-Terme spricht man vom *normalen Zee-*

manneffekt, weil die Aufspaltung ohne den früher noch unbekannten Elektronenspin erklärt werden kann. Dementsprechend bezeichnet man alle anderen Aufspaltungen mit Beteiligung des Elektronenspins als *anomalen Zeemanneffekt*.
In sehr starken Magnetfeldern wird die Kopplung an das äußere Feld so stark, daß die innere Kopplung zwischen $\boldsymbol{L}$ und $\boldsymbol{S}$ zerissen wird. Die auf der Spin-Bahn-Kopplung beruhenden Aufspaltungen sind dann klein gegen die Wechselwirkungsenergie im äußeren Feld. Die damit verbundene Veränderung des Termschemas wird als *Paschen-Back-Effekt* bezeichnet. $\boldsymbol{L}$ und $\boldsymbol{S}$ präzedieren dann unabhängig voneinander um die Richtung von $\boldsymbol{B}$. Die Wechselwirkungsenergie $-\boldsymbol{\mu B}$ nimmt die möglichen Werte $(e\hbar B/2m)(M_L + 2M_S)$ an.
Alle Überlegungen dieses Abschnitts sind auch auf das Wasserstoffatom anwendbar, dann ist $L = l$ und $S = s = 1/2$.

4.2.4.7. Einfluß eines äußeren elektrischen Felds. Ähnlich wie ein magnetisches führt auch ein elektrisches Feld zu einer weiteren Aufspaltung der Energieniveaus. Die sich daraus ergebenden Erscheinungen in den Spektren werden nach ihrem Entdecker als *Starkeffekt* bezeichnet. Durch das angelegte Feld wird die Kugelsymmetrie des Systems zerstört; es verbleibt nur die Drehsymmetrie um die Feldrichtung. Die Symmetrie des Problems und demzufolge auch die Klassifizierung der Elektronenzustände ist die gleiche wie bei zweiatomigen Molekülen (↑ 5.3.1.3.3.). Die mit der Kugelsymmetrie des ungestörten Atoms verbundene $(2I + 1)$-fache Entartung der Zustände mit verschiedener Drehimpulseinstellung wird teilweise aufgehoben. Zustände, die sich nur durch das Vorzeichen der Drehimpulskomponente in Feldrichtung unterscheiden, bei denen also im klassischen Bild lediglich der Drehsinn der Bewegung geändert wird, haben jedoch auch im elektrischen Feld gleiche Energie. Bei Vernachlässigung von Spineffekten würde also ein D-Niveau in drei Niveaus mit $M_L = \pm 2, \pm 1$ und 0 aufspalten.
Die Größe der energetischen Aufspaltung wird durch den Mittelwert der Wechselwirkungsenergie $-\boldsymbol{m}_{el}\boldsymbol{E}$ bestimmt. In hinreichend schwachen Feldern könnte dabei das elektrische Dipolmoment $\boldsymbol{m}_{el}$ des ungestörten Atoms eingesetzt werden. Dieses ist aber im allgemeinen Null, weil die Verteilung der Elektronendichte symmetrisch ist gegenüber einer Spiegelung am Atomkern, so daß der Ladungsschwerpunkt der negativen Elektronen genau im positiv geladenen Kern liegt. Daher gibt es keine feldproportionale Aufspaltung, sondern erst das vom Feld durch Störung der elektronischen Struktur induzierte, feldproportionale Dipolmoment führt zu einer mittleren Wechselwirkungsenergie, die quadratisch von der elektrischen Feldstärke abhängt.
Eine Sonderstellung nimmt das *Wasserstoffatom* ein. Durch die zufällige Entartung (↑ 4.2.2.3.) von Zuständen mit verschiedenen Drehimpulsen können z. B. durch Mischung von s- und p- oder von p- und d-Zuständen in allen angeregten Zuständen ohne Erhöhung der Energie solche Wellenfunktionen gebildet werden, zu denen nichtspiegelsymmetrische Elektronenverteilungen gehören. Daher bilden sich bereits in einem beliebig schwachen Feld Zustände mit einem von der angelegten Feldstärke unabhängigen elektrischen Dipolmoment heraus. Dies führt zu einer feldproportionalen Aufspaltung der Niveaus und der zugeordneten Spektrallinien, was zur Unterscheidung von dem allgemeinen quadratischen Effekt als *linearer Starkeffekt* bezeichnet wird.

4.2.4.8. Hyperfeinstruktur. Das Atom ist ein aus Elektronenhülle und Kern bestehendes Ganzes. Nur die Tatsache, daß bei den üblichen Energien der Hüllenphysik innere Anregungen des Kerns nicht möglich sind, erlaubt es, den Kern als ein nahezu strukturloses Zentrum des Atoms zu behandeln. Gibt es aber irgendwelche mit dem Kern verbundenen Anregungsmöglichkeiten im für die Hülle charakteristischen Energiebereich, so müssen die zugehörigen Energien im Energiespektrum des Atoms berücksichtigt werden.
Alle Kerne mit ungerader und einige wenige mit gerader Nukleonenzahl haben einen von Null verschiedenen Gesamtdrehimpuls, den Kernspin $\boldsymbol{i}$ (↑ 3.3.2.4.), dessen Quadrat wie bei jedem Drehimpuls gequantelt ist und nur die Werte $i(i + 1)\hbar^2$ haben kann. Mit dem Kernspin ist ein magnetisches Moment des Kerns verknüpft, das wegen der verglichen mit der Elektronenmasse wesentlich größeren Masse der Nukleonen um drei Größenordnungen kleiner als die elektronischen Momente der Hülle ist. Die den Kern „umkreisenden" Elektronen, genauer diejenigen in nichtabgeschlossenen Niveaus, erzeugen am Kernort ein Magnetfeld $\boldsymbol{B}_H$, das proportional zum Gesamtdrehimpuls $\boldsymbol{I}$ der Hüllenelektronen ist. Die magnetische Wechselwirkungsenergie $-\boldsymbol{\mu}_K\boldsymbol{B}_H$ ist daher wegen $\boldsymbol{\mu}_K \sim \boldsymbol{i}$ proportional zum Skalarprodukt $\boldsymbol{iI}$ von Kern- und Hüllendrehimpuls und somit von der Orientierung beider Vektoren zueinander abhängig. Die magnetische Kopplung beider Drehimpulse ist ähnlich wie bei der Spin-Bahn-Kopplung der Hüllenelektronen (↑ 4.2.2.6., 4.2.4.4.), aber wegen der Kleinheit der Kernmomente wesentlich schwächer.
Kernspin und Hüllendrehimpuls addieren sich vektoriell zum Atomspin. Sie können sich daher nur so zueinander orientieren, daß das Quadrat des Atomspins entsprechend $\hbar^2F(F + 1)$ quantisiert ist. Die Quantenzahl durchläuft nach den

allgemeinen Regeln mit der Differenz $\Delta F = 1$ benachbarter Zahlen alle Werte zwischen $|I - i|$ und $I + i$. Das sind $2i + 1$ Möglichkeiten für $i \leqq I$ oder $2I + 1$ für $i \geqq I$ mit verschiedenen magnetischen Kopplungsenergien. Auf diese Weise entsteht analog zur Feinstruktur eine Hyperfeinmultiplizität aller Energieniveaus; ebenso wie dort gilt die Landésche Intervallregel (↑ 4.2.4.4.). Die Hyperfeinstrukturaufspaltungen liegen im Energiebereich 10^{-5} eV. Die entsprechende Hyperfeinstruktur der Spektrallinien (↑ 4.3.3.3.2.) kann mit geeigneten Strahlungsquellen, bei denen durch geringe Dichte und niedrige Temperatur der leuchtenden Gase die Linienverbreiterung weitgehend unterdrückt wird, und hochauflösenden Spektralgeräten nachgewiesen werden.
Die Energien der Hyperfeinaufspaltung, also die Energiedifferenzen zwischen verschiedenen Einstellungen des Kernspins, sind so klein, daß im thermischen Gleichgewicht auch bei nahezu vollständig ausgerichteten Hüllendrehimpulsen die verschiedenen Orientierungen des Kernspins bei einer Gesamtheit vieler Atome etwa gleichwahrscheinlich sind. Erst bei extrem tiefen Temperaturen im Bereich 10^{-2} K kann eine räumliche Ausrichtung der Kernspins erreicht werden.

4.2.4.9. Isotopieverschiebungen. Das äußere Erscheinungsbild einer Hyperfeinstrukturaufspaltung der beobachteten Spektrallinien entsteht auch dadurch, daß die Spektren normalerweise nicht von isotopenreinen Substanzen ermittelt werden. Beispielsweise zeigt das von einem Gemisch aus Ne-20 und Ne-22 ausgesandte Neonspektrum Liniendubletts, wobei das Intensitätsverhältnis der beiden Komponenten mit dem Häufigkeitsverhältnis beider Isotope übereinstimmt. Die Atomkerne beider Isotope haben als gg-Kerne keinen Kernspin und damit auch kein magnetisches Kernmoment (↑ 3.3.2.4.), so daß die darauf beruhenden Hyperfeinstrukturaufspaltungen gar nicht auftreten. Die Isotopieverschiebungen beruhen darauf, daß die Energie der Hüllenelektronen ein wenig von den verschiedenen Eigenschaften der isotopen Kerne beeinflußt wird. Dabei dominiert bei leichten Kernen der mit der *Mitbewegung des Kerns* verbundene Einfluß der verschiedenen Kernmassen, der bei wasserstoffähnlichen Systemen (↑ 4.2.2.5., 4.3.3.2.2.) exakt berücksichtigt werden kann. Bei schweren Kernen überwiegt der *Volumeneffekt*. Dabei reagieren besonders die s-Elektronen, die sich mit wesentlich größerer Wahrscheinlichkeit als andere im Kern aufhalten, auf die mit den endlichen Abmessungen des Kerns verbundenen Unterschiede zwischen dem Coulombpotential einer hypothetischen Punktladung und dem auch im Kernmittelpunkt noch endlichen elektrostatischen Potential im realen Kern. Daher treten auch Unterschiede zwischen Atomen mit isotopen Kernen auf, da durch die unterschiedlichen Neutronenzahlen verschiedene Kernradien, also auch verschiedene Ladungsverteilungen im Kern verursacht werden.
Anomal große Volumeneffekte treten als Folge starker Kerndeformationen auf.

4.3. Anregung von Atomen und Atomspektren

4.3.1. Stoßprozesse

Der einzige stabile Zustand eines Atoms ist der energetisch am tiefsten liegende. Um ein Atom aus diesem Grundzustand anzuregen, muß Energie zugeführt werden. Das kann durch Einstrahlung von Licht (↑ 4.1.3.1.) oder durch *inelastische Stöße* (↑ 2.1.3.7.) von Atomen mit anderen Teilchen erfolgen. In vielen Lichtquellen, z. B. Gasentladungsröhre, Lichtbogen und Funkenstrecke, ist dies, trotz der sehr unterschiedlichen technischen Realisierung, der wesentliche Elementarprozeß. Der physikalische Unterschied besteht lediglich in der Art, sowie der Energie- und Richtungsverteilung der verwendeten Stoßpartner. Es ist deshalb von grundsätzlichem physikalischem Interesse, diesen Anregungsmechanismus zu untersuchen. Neben Schlußfolgerungen für optimale technische Parameter folgen gleichzeitig Aussagen über Struktur und Eigenschaften atomarer Systeme, die bei alleiniger Beobachtung ihrer Emissionsspektren nicht zugänglich sind, da diese nicht die Dynamik der Anregung erfassen.

4.3.1.1. Charakteristika für Stoßprozesse. Die experimentelle Anordnung für Atomstoßprozesse besteht so wie bei ähnlichen Untersuchungen in der Kern- und Elementarteilchenphysik aus den drei wesentlichen Teilen Beschleuniger, Target und Nachweissystem. Unter einem *Beschleuniger* versteht man eine Anordnung, die einen Teilchenstrahl einer bestimmten Teilchensorte, also Elektronen, Protonen, α-Teilchen oder Atome und Moleküle, mit einem definierten Ionisierungsgrad und einer festen Energie (hier von einigen eV bis zu einigen 100 MeV) liefert. Durch den Beschleuniger sind also Art, Energie und Richtung der stoßenden Teilchen oder auch Inzidenzteilchen festgelegt. Der eigentliche Stoßprozeß findet im *Target* [*target* engl., Zielscheibe] statt, wo entweder in einer dünnen Folie oder in einem Gasstrahl die Atome, an denen der Stoß erfolgen soll, chemisch rein konzentriert sind. Der Nachweis der Stoßprodukte erfolgt mit *Detektoren*, die unter einem festen Winkel Θ zum Primärstrahl aufgestellt sind. Sie unterscheiden die Stoßprodukte nach Art, Ladungszustand und Energie. Manchmal werden auch die bei einer Anregung

entstandenen Lichtquanten oder Elektronen nachgewiesen. Gegebenenfalls besteht die Nachweisanordnung aus mehreren Detektoren, die zur Sicherung des Nachweises von Teilchen aus der gleichen Elementarreaktion nur gleichzeitig registrieren, d. h. in Koinzidenz geschaltet sind.

Für die Untersuchung von Reaktionen, bei denen im Endzustand wieder nur zwei Teilchen vorliegen, genügt *ein* Detektor. Meist wird für verschiedene Ablenkwinkel Θ die Zahl der abgelenkten Inzidenzteilchen pro Zeiteinheit in Abhängigkeit von ihrer Energie registriert. Da die kinetische Energie der Targetteilchen vor dem Stoß im allgemeinen vernachlässigbar klein ist, ist die Energie der abgelenkten Teilchen höchstens gleich der Inzidenzenergie. Es werden Häufungen von nachgewiesenen Teilchen bei ganz bestimmten Energiewerten beobachtet, die allerdings vom Winkel Θ abhängen. Die Analyse wird dadurch erschwert, daß die beobachteten Energieverluste nicht nur durch inelastische Prozesse, d. h. unter Übergang von Bewegungsenergie in innere Anregungsenergie, sondern auch durch Übertragung kinetischer Energie an den Stoßpartner entstehen. Diese Komplikation verschwindet, wenn man die Meßgrößen in das sogenannte Schwerpunktsystem umrechnet, also in dasjenige Bezugssystem, in dem der Schwerpunkt (↑ 2.1.3.3.) beider Stoßpartner ruht. In diesem System fordert der Impulssatz, daß beide Partner sowohl vor als auch nach dem Stoß betragsmäßig gleich große, aber entgegengesetzt gerichtete Impulse haben. Ist der Stoß elastisch, so sind die Impulsbeträge vor und nach dem Stoß außerdem gleich groß, also bleibt insbesondere die Energie des Stoßteilchens unverändert, und zwar für alle Streuwinkel. Ist der Stoß inelastisch, so zeigt das stoßende Teilchen ebenso wie das gestoßene einen ebenfalls winkelunabhängigen Energieverlust, der jeweils umgekehrt proportional zur Masse des Teilchens ist. Daher kann aus der gemessenen Energie des stoßenden Teilchens sofort die Energie $\Delta E_s = E_s - E_s'$ ermittelt werden, die insgesamt beim Stoß in Anregungsenergie überführt wird. Jedem Wert von ΔE_s ist eine bestimmte Anregungsenergie eines der beiden Stoßpartner zugeordnet. Falls ein optischer Übergang in den Grundzustand erlaubt ist, so kann dieser unabhängig über $\Delta E = h\nu$ aus der Frequenz des ausgestrahlten Lichts ermittelt werden. Ab einem bestimmten Wert von ΔE treten keine diskreten Übergänge mehr auf. Dies entspricht der Ionisierung eines Stoßpartners. Die Ionisierungsenergie läßt sich unabhängig davon durch den Schwellwert der Energie bestimmen, ab der bei einem Stoß freie Elektronen beobachtet werden können.

Über die Lage der Anregungsniveaus hinaus lassen sich aus der relativen Häufigkeit der Stöße weitergehende Informationen gewinnen. Die entsprechende physikalische Größe ist der *Wirkungsquerschnitt* (↑ 2.4.2.4.), den man als Verhältnis der Zahl von gestreuten Atomen pro Zeiteinheit und der Teilchenstromdichte des Primärstrahls bildet. Um von der Größe der Detektorfläche und dem Abstand Target – Detektor unabhängig zu werden, bezieht man diesen Wirkungsquerschnitt noch auf den Raumwinkel $d\Omega$, unter dem der Detektor vom Target aus gesehen wird. Man bezeichnet diese Größe als differentiellen Wirkungsquerschnitt $d\sigma/d\Omega$ der entsprechenden Reaktion. In der klassischen Theorie entspricht $d\sigma$ einem schmalen Kreisring auf der Streuscheibe, d. h. dem Gebiet des Stoßparameters s, der zu einer Ablenkung in den Raumwinkel $d\Omega$ führt. Dabei wird als Stoßparameter der Abstand der Anflugsgeraden zum Stoßzentrum bezeichnet. Die Messung des differentiellen Wirkungsquerschnitts liefert also eine Beziehung zwischen dem Ablenkwinkel und diesem Stoßparameter s. Da verschiedene Stoßparameter s auch zu verschiedenen klassischen Bahnen des Inzidenzteilchens relativ zum Targetatom führen, ist anschaulich klar, daß man aus einer solchen Relation eine tiefere Einsicht in die atomaren Strukturen von beiden Stoßpartnern gewinnen kann. Auf eine quantenmechanische Beschreibung des Stoßvorgangs soll hier der Einfachheit halber verzichtet werden.

4.3.1.2. Stoß einfacher Teilchen mit Atomen. Beim Stoß zwischen einem Elektron, Proton oder α-Teilchen und einem Atom können jene als strukturlos, gewissermaßen als punktförmig betrachtet werden. Wie RUTHERFORD und seine Schüler GEIGER und MARSDEN bereits 1908/11 bei der Analyse der Streuung von α-Teilchen an Atomen verschiedener Substanzen zeigen konnten (↑ 1.1.4.1.), ergibt sich im Schwerpunktsystem für den differentiellen Wirkungsquerschnitt der elastischen Streuung in weiten Grenzen die Beziehung $d\sigma/d\Omega = (b/4)^2/(\sin\vartheta/2)^4$. Dem entspricht die Abhängigkeit $\operatorname{tg}\vartheta/2 = b/2s$ des Ablenkwinkels vom Stoßparameter. Die hier auftretende Länge b wird Streudurchmesser genannt, weil alle Stoßvorgänge mit $s \leqq b/2$, d. h. innerhalb eines Kreises mit dem Durchmesser b, zu Ablenkungen $\vartheta \geqq \pi/2$, d. h. zu Rückwärtsstreuung, führen. Dabei ergab sich zwischen b und den Ordnungs- bzw. Massenzahlen der Stoßpartner sowie der Energie E des mit 1 indizierten stoßenden Teilchens $b = Z_1 Z_2 e^2 (M_1 + M_2)/(4\pi\varepsilon_0 E M_2)$. Diese beobachteten Gesetzmäßigkeiten folgen aus der Annahme, daß sich die punktförmig konzentrierten Ladungen $Z_1 e$ und $Z_2 e$ der beiden Stoßpartner elektrostatisch abstoßen. Ein Einfluß der tatsächlichen Ladungsverteilung im Atomkern ist bei den verwendeten

Energien von einigen MeV nicht zu erwarten, weil hier der Stoßdurchmesser b bei einigen 10^{-14} m noch um etwa eine Größenordnung über der Ausdehnung des Kerns liegt. Erst bei Erhöhung der Energien um ein bis zwei Größenordnungen ist für kleine Stoßparameter, d. h. in der Rückwärtsstreuung, mit einem Eindringen des α-Teilchens in den Kern zu rechnen. Die beobachteten Effekte konnten experimentell zur Bestimmung der Kernradien (↑ 3.2.2.1.) herangezogen werden.

Die Ladungsverteilung der Hüllenelektronen, die ja eine anziehende Kraft auf das α-Teilchen ausübt, spielte in den damaligen Experimenten keine Rolle, weil der Stoßdurchmesser $b \approx 10^{-14}$ m um drei bis vier Größenordnungen kleiner als die charakteristische Hüllenausdehnung $a \approx 10^{-11}$ bis 10^{-10} m ist. Stoßvorgänge mit merklichen Ablenkungen ϑ von einigen Grad kommen nach dem gefundenen Stoßgesetz $s = b/2\,(\text{tg}\,\vartheta/2) \approx b/\vartheta$ (ϑ hier im Bogenmaß einige 10^{-2}) nur bei Stoßparametern $s \ll a$ vor, weshalb der Einfluß des Kerns noch immer der bestimmende sein muß. Nur für langsamere Teilchen, d. h. kleinerer Energie und deshalb größeren Stoßdurchmesser b, tritt noch eine merkliche Ablenkung durch den Kern für Flugbahnen mit Stoßparametern $s \approx a$ auf. Die anziehende Wirkung der Hüllenelektronen müßte in solchen Fällen zu einer beobachtbaren Verringerung der Ablenkung des Projektils führen, weshalb der differentielle Wirkungsquerschnitt für kleine Energien und Winkel unter dem nach dem Rutherfordmodell zu erwartenden Wert liegen sollte. Solche Effekte bieten die Chance, eine direkte Aussage über die Ladungsverteilung der Elektronen in der Atomhülle zu erhalten, da man die Kräfte zwischen Projektil und Hüllenelektronen theoretisch gut kennt und nur noch ihre Ladungsverteilung in den Stoßvorgang als Unbekannte eingeht. Tatsächlich liegen die Meßergebnisse für $s \approx a$ unter den Werten, die nach dem Rutherfordmodell zu erwarten wären, wie Abb. 4.3.1-1 zeigt. Schon ein einfaches statistisches Atommodell (↑ 4.2.3.5.), das für alle Atome einen ähnlichen Verlauf des radialen Dichteverlaufs annimmt, liefert im Vergleich mit dem Experiment, hier mit der Streuung von He^{+} an Kryptonatomen, im Mittel eine recht gute Beschreibung der abschirmenden Wirkung der Hülle. Allerdings oszillieren die Meßwerte um diese mittlere Kurve z. T. recht erheblich, was auf einen nichtmonotonen radialen Verlauf der Elektronendichte schließen läßt. Ein Vergleich mit einer Schalenmodellrechnung zeigt, daß man diese Oszillationen tatsächlich mit den Lagen der Maxima der Elektronendichten der K-, L-, M-Schalen in Verbindung bringen kann. Man hat so mit der Messung der Streuung von α-Teilchen, aber auch der von Protonen und Elektronen, eine direkte Möglichkeit zur Ausmessung der Ladungsverteilung im gesamten Atom. Diese Information ist wesentlich direkter als die Aussagen, die man aus Experimenten mit Licht- oder Röntgenquanten gewinnen kann, weil hier nur bestimmte Mittelwerte der Verteilung, z. B. Dipolmomente, in die beobachtbaren Größen wie Brechungs- oder Absorptionsindex eingehen.

Abb. 4.3.1-1 Elastischer differentieller Wirkungsquerschnitt für die Streuung von He^{+}-Ionen an Krypton; Vergleich verschiedener Modellrechnungen mit experimentellen Messungen

Der inelastische Stoß wurde, wenn auch nicht in einem direkten Streuexperiment, zuerst 1913 von FRANK und HERTZ beim Stoß von Elektronen an Hg-Dampfatomen nachgewiesen. In dem historisch bedeutsamen Frank-Hertz-Versuch (↑ 1.1.4.4.) wurde erstmals die Existenz diskreter Anregungsenergien atomarer Systeme direkt experimentell nachgewiesen.

4.3.1.3. Atom-Atom-Stöße. Auch beim Stoß von zwei komplexen atomaren Systemen, z. B. eines schweren Ions mit einem Gasatom, spielt das Coulombpotential die entscheidende Rolle für die Ablenkung. Für kleinere Energien und Winkel ist die abschirmende Wirkung beider Elektronenhüllen zu berücksichtigen. Eine Diskussion dieser Doppelabschirmung vom Standpunkt der Wechselwirkung der atomaren Ladungsverteilungen isolierter Atome ist aber schon deshalb nicht mehr möglich, weil infolge des Pauliprinzips ein abstoßend wirkender Term im Wechselwirkungspotential durch die Überlappung beider Elektronendichten zu erwarten ist, auch wenn man von einer Deformation absieht. Dies kann durch die Analyse der elastischen Streuung bestätigt werden. Ein

direkter Nachweis von Deformationen konnte durch die Untersuchung des emittierten Röntgenspektrums bei der Streuung angeregter Nb-Atome an Nb erbracht werden (Abb. 4.3.1-2). Während im Spektrum eines isolierten Nb-Atoms nur die beiden K-Linien auftreten, die einem Einfang eines Elektrons in die unbesetzte K-Schale entsprechen (↑ 4.3.3.2.3.), treten beim Stoß von Nb an Nb Röntgenquanten auf, deren Energie wesentlich größer als die der K-Strahlung ist. Es konnte gezeigt werden, daß es sich um die Emission von Röntgenstrahlung handelt, die während des Stoßvorgangs von dem gebildeten Quasimolekül ausgesandt wird.

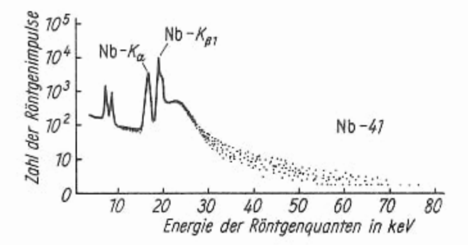


Abb. 4.3.1-2 Röntgenspektrum gemessen beim Beschuß eines dicken Nb-Targets mit 65 MeV Nb-Ionen; die energetisch höchsten Röntgenübergänge des isolierten Nb-Atoms sind gekennzeichnet

Dabei bildet sich bei kleiner Relativgeschwindigkeit der Nb-Atome nahezu für jeden Abstand ein Quasimolekülzustand. Im Grenzfall der Berührung beider Atomkerne entspricht das gebildete Quasimolekül einem Atom mit der Kernladung 2 Z. Wegen der Proportionalität der K-Strahlungsenergie zu Z^2 dürften K-Strahlungen des Quasimoleküls nur bis etwa zum vierfachen Energiewert des Einzelatoms auftreten, was auch am gezeigten Spektrum ersichtlich ist. Dieses Beispiel zeigt, daß man bei Stoßvorgängen atomare Systeme beobachten kann, die in der Natur nicht stabil vorkommen. Damit lassen sich die Gültigkeitsgrenzen bekannter theoretischer Modelle bis zu extremen Bedingungen testen.

Von *Stößen zweiter Art* spricht man, wenn während des Stoßvorgangs die Anregungsenergie eines Atoms auf das andere übertragen wird. Als ein möglicher Mechanismus kann hier wieder die Ausbildung eines angeregten Quasimoleküls angesehen werden, das bei seiner Bildung bzw. seinem Zerfall seine Anregung von dem einen Atom erhält bzw. an das andere abgibt. Eine Anwendung finden solche Prozesse z. B. im He–Ne-Gaslaser. Hier erfolgt durch Elektronenstoß primär eine bevorzugte Anregung der He-Atome, die diese durch Stöße zweiter Art auf die Ne-Atome übertragen. Das führt zu einer Besetzung des oberen Laserniveaus im Ne-Atom und so zur Ausbildung des eigentlichen Lasereffekts im Ne-System (↑ 4.3.3.).

Eine Möglichkeit, den Stoßvorgang zwischen Atomen zur Anregung extrem hoher Anregungszustände auch von Ionen zu nutzen, bietet die sogenannte *beam-foil-Spektroskopie* [*beam-foil* engl., Strahl-Folie]. Die zu untersuchenden Ionen werden nach Beschleunigung auf ein dünnes Target, gewöhnlich eine Kohlenstofffolie, geschossen. Dabei werden die Ionen angeregt und verlieren oder gewinnen durch Austausch mit den Targetatomen Elektronen. Teilt man den Strahl, z. B. durch ein elektrisches Ablenksystem, in die verschiedenen Ladungszustände auf, kann man die Anregung der verschiedenen möglichen Ionen eines Atoms getrennt untersuchen. Durch den Nachweis der emittierten Lichtquanten in verschiedenen Abständen lassen sich Aussagen über die Lebensdauer der Anregungszustände und die Aufeinanderfolge der verschiedenen Übergänge machen, verschiedene Abstände von der Streufolie entsprechen verschiedenen Zeiten nach der Anregung. Mit dieser Methode konnten sehr eindeutige und besonders für die Astronomie sehr heißer Sterne wichtige Grunddaten gesammelt werden.

4.3.2. Atomspektralanalyse

Die Atomspektralanalyse wurde in ihren grundlegenden Zügen bereits vor über 100 Jahren von Kirchhoff und Bunsen entwickelt, also lange bevor die theoretischen Zusammenhänge zwischen Atomstruktur und -spektrum bekannt waren. Wegen ihrer hohen Nachweisempfindlichkeit und Genauigkeit hat die Atomspektralanalyse seither eine ständig wachsende Bedeutung für die physikalische, chemische und biologische Forschung sowie für Anwendungen in Industrie, Landwirtschaft, Medizin und Umweltschutz erhalten. Jedes Element vermag bei geeigneter Anregung eine Vielzahl von Spektrallinien zu emittieren, die durch *Wellenlänge*, *Intensität*, *spektrales Profil* und *Polarisation* gekennzeichnet sind (↑ 4.1.1., 4.1.3.). Die Gesamtheit der ausgestrahlten Spektrallinien, also die Abhängigkeit der ausgestrahlten Intensität von der Wellenlänge bzw. Frequenz stellt das *Emissionsspektrum* dar. Bei den Wellenlängen vieler Emissionslinien können die Atome bei Einstrahlung von Licht umgekehrt auch eine Absorption zeigen (↑ 4.1.3.1.). die Abhängigkeit der Intensität des durch die Atomprobe hindurchtretenden Lichts von der Wellenlänge wird dementsprechend als *Absorptionsspektrum* bezeichnet. Solche Absorp-

tionen wurden zuerst durch FRAUNHOFER (1813) als dunkle Linien im kontinuierlichen Sonnenspektrum beobachtet (Tafel 67) und von KIRCHHOFF und BUNSEN (1859) aufgrund der Analyse von Emissionslinien gleicher Wellenlänge bestimmten Elementen zugeordnet, die sich in diesem Falle als Gas in der Umgebung der heißen, weißes Licht ausstrahlenden Sonne befinden (s. a. 1.1.4.2.).
Die Wellenlängen und Intensitäten der Spektrallinien hängen unmittelbar mit dem *Aufbau der Elektronenhülle der Atome* (↑ 4.2.3.) zusammen und ihre Messung und theoretische Deutung kann deshalb einen Einblick in die Atomstruktur vermitteln. Die Analyse der Atomspektren hat bei der Entwicklung von physikalisch begründeten Modellen für den Atomaufbau und besonders für die Herausbildung der Quantentheorie eine wichtige Rolle gespielt. Auch gegenwärtig können durch die Registrierung von Spektren mit immer höherer Auflösung, z. B. mittels der Laserspektroskopie (↑ 4.1.4.7.5.), noch neue Informationen über Wechselwirkungen im Atom gewonnen werden.
Die Emissions- und Absorptionsspektren sind so typisch für ein bestimmtes Element, daß sie wie Fingerabdrücke zur Auffindung und Kennzeichnung dieser Elemente dienen können. Diese Methode, die nicht unbedingt die Kenntnis des Zusammenhangs zwischen Aufbau und Spektrum voraussetzt, sondern mittels Spektrenvergleichs durchgeführt werden kann, wird sowohl zur qualitativen als auch quantitativen Analyse erfolgreich eingesetzt.
Bei der *qualitativen Spektralanalyse* wird aus dem Vorhandensein typischer Spektrallinien auf das Vorhandensein des entsprechenden Elements geschlossen, indem die registrierten Wellenlängen mit den Linien in Spektralatlanten verglichen werden. Die Empfindlichkeit der besten Verfahren bei einer *Spurenanalyse* ist so hoch, daß dabei einzelne Verunreinigungsatome in einer Gasprobe nachgewiesen werden können.
Bei der *quantitativen Spektralanalyse* wird aus der Intensität der registrierten typischen Spektrallinie die Konzentration des entsprechenden Elements bestimmt, indem man einen Vergleich mit Spektren von Proben bekannter Konzentration durchführt. Dabei kann Eichung und Auswertung mit sogenannten *Quantometern* vollautomatisch durchgeführt werden.

4.3.2.1. Emissionsspektralanalyse. Die Atome werden in die Gasphase gebracht und durch Energiezufuhr in ein höheres Niveau angeregt, von dem aus sie durch *spontane Emission* (↑ 4.1.3.1.) in niedrigere Niveaus übergehen. Das anschließend emittierte Licht wird spektral untersucht.

4.3.2.1.1. *Anregung der Emissionsspektren.* Die Anregung der Emissionsspektren kann durch Elektronen-, Atom- oder Ionenstöße (↑ 4.3.1.) oder durch Lichtabsorption erfolgen. Die Stoßanregung erreicht man durch Beschuß der Atome mit entsprechenden Teilchenstrahlen oder auch in Glimmentladungen (Tafel 79), in Lichtbögen, in Funkenentladungen, in Flammen sowie in aufgeheiztem, verdampftem Material, das durch Fokussierung von Laserimpulsen auf Festkörperoberflächen entsteht. Die Materialanregung mit fokussierter Laserstrahlung bietet die Möglichkeit einer Mikrospektralanalyse, bei der Untersuchungsbereiche von einigen Mikrometern Durchmesser auf einer Probe mittels eines Mikroskops ausgewählt werden können (Tafeln 20/21).
Sehr scharfe Spektrallinien liefert die Anregung durch *Glimmentladung*, eine Niederdruckentladung zwischen kalten Elektroden. Außer den Atomen und Ionen des Füllgases können dabei auch die ursprünglich im Katodenmaterial enthaltenen Elemente angeregt werden, die infolge von Materialzerstäubung durch den Aufprall von Ionen auf die Katode freigesetzt werden. Sehr bewährt haben sich für diesen Zweck die *Hohlkatodenlampen*, bei denen die Katode als zylindrischer Hohlkörper ausgebildet ist, in den sich bei bestimmten Drücken das katodennahe Glimmlicht der Entladung hineinzieht und den Zerstäubungseffekt auf der entsprechend präparierten Innenfläche hervorruft. Zur Anregung durch Lichtabsorption, also zur Anregung von Fluoreszenzspektren, können Emissionslinien geeigneter Wellenlänge, das mittels Monochromator gefilterte Licht thermischer Quellen oder die Strahlung abstimmbarer Laser verwendet werden. Eine besonders hohe Selektivität der Anregung ($\Delta\nu \approx 10^6$ Hz) kann mit abstimmbaren Farbstofflasern (Tafel 78) erreicht werden.

4.3.2.1.2. *Spektrometer zur Untersuchung von Emissionsspektren.* Sie werden auf der Basis von geeigneten Anregungsanordnungen (↑ 4.3.2.1.1.) sowie von Prismen- oder Gitterspektrografen oder von Interferometern aufgebaut (↑ 4.1.1.1.). Der Nachweis des spektral zerlegten Lichts erfolgt mit Fotoplatten oder fotoelektrischen Empfängern (↑ 4.1.1.3.). Besonders hohe Nachweisempfindlichkeiten und Genauigkeiten werden mit der Photonenzählung erreicht. In Emissionsspektren für laserangeregte Mikroanalysen kann das durch Laserbestrahlung verdampfte Material zusätzlich elektrisch angeregt werden.
Bei Fluoreszenzuntersuchungen kann die Registrierung entweder in Abhängigkeit von der Anregungswellenlänge bei fester Beobachtungswellenlänge oder in Abhängigkeit von der Beobachtungswellenlänge bei fester Anregungswellenlänge erfolgen.

Atom-Fluoreszenzspektroskopie wird außer im ultravioletten Spektralbereich auch im Röntgengebiet betrieben, wo besonders typische Spektrallinien auftreten (↑ 4.3.3.2.3.). Die emittierte Strahlung wird mit Kristallspektrometern (↑ 4.1.1.1.1.) untersucht. Ein großer Vorteil der *Röntgenfluoreszenzspektroskopie* besteht darin, daß die nachzuweisenden Atome in beliebigen Verbindungen und diese in beliebigem Aggregatzustand untersucht werden können, weil die Emissionslinien im Röntgengebiet nur sehr wenig durch Wechselwirkungen der äußeren Hüllenelektronen beeinflußt werden. Dadurch kann auf eine besondere Probenpräparation verzichtet und eine hohe Genauigkeit bei Analysen erzielt werden.

4.3.2.2. Absorptionsspektralanalyse. Die zu untersuchenden Atome werden in die Gasphase gebracht. Gemessen wird die Absorption einer möglichst monochromatischen Lichtstrahlung, deren Wellenlänge mit einer der Absorptionslinien der Atome übereinstimmt.

Die mit weißen Lichtquellen und üblichen *Monochromatoren* erreichbare Frequenzschärfe ist im allgemeinen nicht ausreichend. Als Lichtquelle hoher Monochromasie können z. B. Hohlkatodenlampen (↑ 4.3.2.1.1.) verwendet werden. Da diese Lampen nur eine oder wenige Linien fester Frequenzen ausstrahlen, benötigt man aber für den Nachweis jeden Elements eine angepaßte Hohlkatodenlampe. Günstig können abstimmbare Farbstofflaser mit hoher Einstellgenauigkeit und Monochromasie eingesetzt werden.

Die Absorption wird durch Messung der Lichtintensität vor und hinter der Probe bestimmt. Mittels eines Rechners können Eichspektren und Probenspektren automatisch ausgewertet und Konzentrationen berechnet werden.

Ein sehr empfindliches Meßverfahren für kleine Absorptionen stellt die intracavity-Spektroskopie dar [*intracavity* engl., innerhalb eines Hohlraums oder Resonators]. Dabei wird die Probe mit absorbierenden Atomen im *Laserresonator* untersucht, dessen aktives Medium ein breites Verstärkungsprofil besitzt (↑ 4.1.4.1.). Wenn der Laser in der Nähe der Schwelle betrieben wird, ist seine Ausgangsleistung bei den Wellenlängen, bei denen die Probe eine Absorption zeigt, wesentlich kleiner als in absorptionsfreien Spektralgebieten.

4.3.3. Frequenzen und Intensitäten der Atomspektren, Auswahlregeln

Als in den ersten Jahrzehnten des 20. Jahrhunderts die Quantentheorie entwickelt wurde, existierte die Atomspektralanalyse (↑ 4.3.2.) schon fast ein halbes Jahrhundert. Die Spektren vieler Atome waren schon vermessen worden. Die Spektralanalyse hatte den Nachweis geliefert, daß Atome oder Atomionen *Linienspektren* emittieren oder absorbieren. Die Spektroskopie konnte die Wellenlängen der einzelnen Spektrallinien im Linienspektrum mit sehr großer Genauigkeit bestimmen. Die große Meßgenauigkeit der Atomspektralanalyse war von entscheidender Bedeutung für den quantitativen Vergleich von theoretischen Ergebnissen mit dem empirischen Faktenmaterial. Bei der Entwicklung der Quantentheorie und ihrer Anwendung zur Aufklärung der Struktur der Atomhülle war die Erklärbarkeit von spektroskopischen Meßergebnissen das wichtigste Kriterium für die Richtigkeit der Theorie.

4.3.3.1. Allgemeine Auswahlregeln bei Strahlungsübergängen. Eine Betrachtung von Atomemissions- oder -absorptionsspektren zeigt, daß nicht alle Spektrallinien gleich stark erscheinen. Es kann sogar passieren, daß bei bestimmten Wellenlängen, bei denen gemäß dem Energieniveausystem des Atoms nach der Bohrschen Bedingung (↑ 5.3.2.) Übergänge liegen, keine Spektrallinien beobachtet werden.

Die Stärke der Spektrallinien in der Emissionsspektroskopie ist ein Maß für die Strahlungsleistung des von angeregten Atomen spontan emittierten Lichts. Die Strahlungsleistung L beim Übergang von N_i angeregten Atomen aus dem Zustand i in einen Zustand j ist gegeben durch $L = h\nu_{ij}N_iA_{ij}$; ν_{ij} ist die Frequenz der emittierten Photonen und A_{ij} ist der Einsteinkoeffizient für spontane Emission (↑ 4.1.3.1.).

In der Absorptionsspektroskopie wird so verfahren, daß eine Probe mit der Dicke Δx mit Licht der Leistung L bestrahlt wird. Gemessen wird die Lichtleistungsänderung ΔL infolge Absorption beim Durchgang des Lichts durch die Probe:

$$(1)\quad \Delta L = -Ln_jB_{ji}h\nu_{ij}g(\nu - \nu_{ij})\frac{\Delta x}{c},$$

n_j ist die Anzahldichte der Atome im unteren Energieniveau j, B_{ji} ist der Einsteinkoeffizient für Absorption (↑ 4.1.3.2.), $g(\nu - \nu_{ij})$ ist die Linienformfunktion für den Übergang $j \to i$ (↑ 4.1.4.1.). Die Stärke der Spektrallinien in der Emissions- und Absorptionsspektroskopie hängt demgemäß von den Einsteinkoeffizienten für die entsprechenden Übergänge, von den Übergangsfrequenzen und von den Besetzungszahlen der Atome in den Ausgangszuständen ab. Die Besetzungszahlen werden durch die Anregungsbedingungen bestimmt; im thermodynamischen Gleichgewicht genügen die Besetzungszahlen der Boltzmannverteilung (↑ 2.4.4.7.1.).

Die *Einsteinkoeffizienten* werden unmittelbar durch die Atomeigenschaften bestimmt. Aus der

Quantentheorie folgt für Übergänge, an denen nur ein Elektron beteiligt ist

(2) $$B_{ji} = \frac{2\pi^2 e^2}{3h^2\varepsilon_0} \left| \sum_{m_s} \int \mathrm{d}V \varphi_i^*(r, m_s)\, r \varphi_j(r, m_s) \right|^2$$

(3) $$A_{ij} = \frac{16\pi^3 \nu_{ij}^3 e^2}{3c^3 h \varepsilon_0} \times \left| \sum_{m_s} \int \mathrm{d}V \varphi_i^*(r, m_s)\, r \varphi_j(r, m_s) \right|^2$$

B_{ji} hängt bis auf Naturkonstanten nur vom Betragsquadrat des sogenannten Übergangsmoments $\sum_{m_s} \int \mathrm{d}V \varphi_i^* r \varphi_j$ ab. A_{ij} ist zusätzlich noch der 3. Potenz der Übergangsfrequenz ν_{ij} proportional. Die Größe des Übergangsmoments bestimmt wesentlich die Intensitäten der Spektrallinien für die entsprechenden Übergänge. φ_i und φ_j sind die von den Raumkoordinaten und der Spinvariablen m_s abhängigen Wellenfunktionen des Anfangs- und Endzustands (s. a. 2.5.5.4.). Der Ortsvektor r geht vom Atomkern aus.

Die Rechnungen zeigen, daß die Übergangsmomente nur bei Übergängen zwischen bestimmten Anfangs- und Endzuständen von Null verschieden sind. Ein Übergang mit einem Übergangsmoment ungleich Null wird als *optisch erlaubter Übergang* bezeichnet. Er kann im Atom durch Absorption oder Emission von Strahlung vonstatten gehen. Ein Übergang mit einem Übergangsmoment gleich Null wird als *optisch verbotener Übergang* bezeichnet. Nur die optisch erlaubten Übergänge führen zu Spektrallinien.

Die Berechnung der Übergangsmomente ist für den Fall eines Elektrons in einem Zentralkraftfeld und bei vernachlässigbarer Spin-Bahn-Kopplung (↑ 4.2.2.6.) noch relativ einfach. Die spinabhängigen Wellenfunktionen beschreiben dann Zustände mit definierten Spineinstellungen m_{si} und m_{sj} des Elektrons. Das Übergangsmoment ist dann nur von Null verschieden, wenn die Spinkomponenten von Anfangs- und Endzustand gleich sind, d. h. $m_{si} = m_{sj}$. Ist diese Bedingung erfüllt, so ist das Übergangsmoment durch $\int \mathrm{d}V \varphi_i(r)\, r \varphi_j(r)$ bestimmt, φ_i und φ_j sind die nur noch von den räumlichen Koordinaten abhängigen Wellenfunktionen des Anfangs- bzw. Endzustands. Dieses Integral ist nur dann ungleich Null, wenn zwischen den Bahndrehimpulsquantenzahlen l und magnetischen Quantenzahlen m der Anfangs- und Endzustände die Bedingungen $\Delta l = l_i - l_j = \pm 1$ und $\Delta m = m_i - m_j = 0$ oder ± 1 erfüllt sind. Solche Beziehungen werden als *Auswahlregeln für Strahlungsübergänge* bezeichnet. Eine Auswahlregel für die Hauptquantenzahl n besteht nicht.

Im allgemeinen Fall eines Vielelektronenatoms gelten nur Auswahlregeln für die Quantenzahlen I und M für den Gesamtdrehimpuls und seine z-Komponente, und zwar $\Delta I = 0$ oder ± 1, $\Delta M = 0$ oder ± 1; ($I = 0 \rightarrow I = 0$ verboten). Sie formulieren einfach die Drehimpulserhaltung, da das Lichtquant einen Drehimpuls mit der Betragsquantenzahl 1 besitzt. Dabei wird die Auswahlregel für die Drehimpulskomponente M nur dann wichtig, wenn durch ein äußeres magnetisches oder auch elektrisches Feld die Zustände mit verschiedenen M-Werten auch verschiedene Energien haben.

Falls man die Spin-Bahn-Kopplung vernachlässigen kann, werden die Atomzustände zusätzlich durch die Quantenzahlen L und S für den gesamten Bahn- bzw. Spindrehimpuls der Hüllenelektronen (↑ 4.2.4.2.) gekennzeichnet. Dann gelten zusätzlich die Auswahlregeln $\Delta L = 0$ oder ± 1, $\Delta S = 0$. Entsprechend dem sogenannten Interkombinationsverbot $\Delta S = 0$ gibt es keine Strahlungsübergänge zwischen Elektronenzuständen mit verschiedenen S-Werten.

Wird die zur Hyperfeinstruktur (↑ 4.2.4.8.) führende Kopplung zwischen Gesamtdrehimpuls der Elektronenhülle I und dem Kernspin i berücksichtigt, kommt noch die Auswahlregel für die Quantenzahl des Gesamtdrehimpulses von Kern und Hülle, dem sogenannten *Atomspin* F, hinzu: $\Delta F = 0$, ± 1 ($F = 0 \rightarrow F = 0$ verboten).

Außer dem strikten Verbot von 0-0-Übergängen gelten die einzelnen Auswahlregeln nicht absolut streng, sondern können durch Störungen infolge von äußeren Feldern oder durch inneratomare Wechselwirkungen, die bei der Ableitung der Regeln vernachlässigt wurden, verletzt werden.

4.3.3.2. Serienstruktur der Atomspektren. Die Linienspektren von Atomen mit vielen Hüllenelektronen sind sehr kompliziert (Tafel 67). Das einfachste Linienspektrum ist das Wasserstoffspektrum, deshalb wurden zuerst in ihm gesetzmäßige Zusammenhänge zwischen den Wellenlängen der einzelnen Spektrallinien entdeckt. In Tafel 67 ist das Wasserstoffatomspektrum für den sichtbaren Spektralbereich dargestellt. Es ist deutlich zu erkennen, daß die Spektrallinien mit kürzer werdenden Wellenlängen immer enger zusammenrücken. Balmer entdeckte zwischen den Wellenlängen der Spektrallinien des sichtbaren Wasserstoffatomspektrums einen gesetzmäßigen Zusammenhang und gab 1885 als erster eine empirisch gewonnene Formel zur Berechnung der Wellenlängen der Linien an.

Eine Folge von Spektrallinien mit gesetzmäßigem Zusammenhang zwischen den Wellenlängen der Linien wird als *Spektralserie* bezeichnet. Formelmäßig wird dieser Zusammenhang durch das *Seriengesetz* ausgedrückt. Später wurden im ultravioletten und infraroten Bereich

weitere Spektralserien des Wasserstoffspektrums gefunden (↑ 4.3.3.2.1.). Auch in den Spektren anderer Atome wurden gesetzmäßige Zusammenhänge zwischen den Wellenlängen der Spektrallinien entdeckt, und die Seriengesetze für die Spektralserien aufgestellt (s. a. 4.3.3.2.3.). Anhand der verschiedenen Linienspektren von neutralen Atomen konnte RYDBERG zeigen, daß sich die Wellenlängen λ der Spektrallinien häufig durch ein Seriengesetz der Form

$$\frac{1}{\lambda} = \frac{R}{(m+a)^2} - \frac{R}{(n+b)^2} \tag{4}$$

beschreiben lassen. Dabei sind λ die Wellenlängen der Spektrallinien, a und b charakteristische Konstanten für eine Spektralserie, m und n natürliche Zahlen, wobei m innerhalb einer Serie konstant bleibt, während n alle natürlichen Zahlen $> m$ durchläuft. Die Größe R ist atomspezifisch und wird als Rydbergkonstante bezeichnet.

Mit wachsendem n konvergieren die Wellenlängen der Spektrallinien gegen eine untere Grenzwellenlänge. Man bezeichnet diese als *Seriengrenze* der Spektralserie, an die sich das sogenannte *Seriengrenzkontinuum* anschließt. Mit Hilfe der Quantentheorie gelang es, die Seriengesetze zu erklären und die empirischen Konstanten physikalisch zu deuten.

4.3.3.2.1. *Wasserstoffspektrum.* Neben der Balmerserie wurden im infraroten und ultravioletten Spektralbereich weitere Spektralserien für Wasserstoff gefunden. Die zu den einzelnen Spektrallinien gehörigen Atomübergänge sind in Abb. 4.3.3-1 eingezeichnet. Die verschiedenen Spektrallinien einer Spektralserie entstehen durch Übergänge zwischen gleichen unteren Energieniveaus und verschiedenen höheren

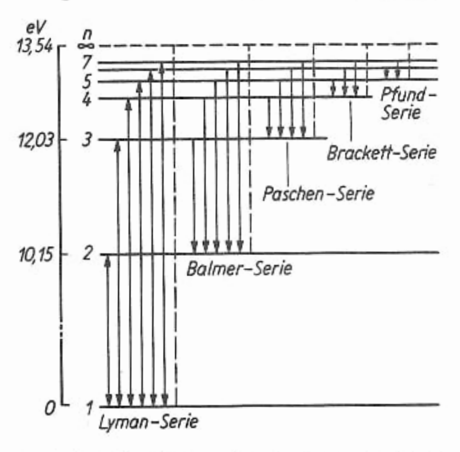


Abb. 4.3.3-1 Niveauschema und Spektralserien des H-Atoms

Niveaus. Da zu jedem durch die Hauptquantenzahl n gekennzeichneten Niveau Elektronenzustände mit Drehimpulsquantenzahlen $l = 0, 1, \ldots, (n-1)$ gehören (↑ 4.2.2.3.), wird die Auswahlregel $\Delta l = \pm 1$ nicht wirksam. Es sind Übergänge zwischen beliebigen n-Werten möglich. Im folgenden sind die Seriengesetze der bekanntesten Spektralserien des H-Atoms zusammengestellt:

$$\frac{1}{\lambda} = R_H\left(1 - \frac{1}{n^2}\right) \quad n = 2, 3, \ldots \quad \text{Lymanserie}$$

$$\frac{1}{\lambda} = R_H\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \quad n = 3, 4, \ldots \quad \text{Balmerserie}$$

$$\frac{1}{\lambda} = R_H\left(\frac{1}{3^2} - \frac{1}{n^2}\right) \quad n = 4, 5, \ldots \quad \text{Paschenserie}$$

$$\frac{1}{\lambda} = R_H\left(\frac{1}{4^2} - \frac{1}{n^2}\right) \quad n = 5, 6, \ldots \quad \text{Brackettserie}$$

$$\frac{1}{\lambda} = R_H\left(\frac{1}{5^2} - \frac{1}{n^2}\right) \quad n = 6, 7, \ldots \quad \text{Pfundserie}$$

Die Spektralserien wurden nach ihren Entdeckern benannt. Anhand der Spektrallinien konnte die Rydbergkonstante R_H für das H-Atom sehr genau bestimmt werden. Laserspektroskopische Methoden ergaben $R_H = 109\,677{,}575\ \text{cm}^{-1}$. Im Energiemaß ist R_H gleich der Ionisierungsenergie des H-Atoms oder gleich der Energie der absorbierten bzw. emittierten Photonen bei der Seriengrenzwellenlänge $\lambda_G = 91{,}1$ nm der Lymanserie. Die Ionisierungsenergie des H-Atoms beträgt 13,53 eV.

4.3.3.2.2. *Wasserstoffähnliche Spektren.* Schon die Bohrsche Theorie konnte zeigen, daß für Einelektronenatome zwischen den Energien $E(n)$ in den verschiedenen Atomniveaus und der Kernladung Z die Beziehung $E(n) = -hc \times (Z^2R/n^2)$ besteht. Danach müßten die Wellenlängen der Spektrallinien von He^+- oder Li^{++}-Ionen bis auf einen größeren Vorfaktor den Seriengesetzen des H-Atoms genügen. Es folgt

$$\frac{1}{\lambda} = 4R_H\left(\frac{1}{m^2} - \frac{1}{n^2}\right) \quad \text{für } He^+\text{-Ionen und}$$

$$\frac{1}{\lambda} = 9R_H\left(\frac{1}{m^2} - \frac{1}{n^2}\right) \quad \text{für } Li^{++}\text{-Ionen.}$$

Spektroskopische Untersuchungen zeigten, daß die Wellenlängen der Spektrallinien von He^+- und Li^{++}-Ionen bis auf einen kleinen Korrekturfaktor nach diesen Seriengesetzen berechnet werden können. Die Spektren werden deshalb als *wasserstoffähnlich* bezeichnet. Der Korrekturfaktor ist auf die unterschiedlichen Massen der Atomkerne von Wasserstoff-, Helium- und Lithiumatomen zurückzuführen. Die Beiträge der *Mitbewegung des Atomkerns* zur Rydberg-

konstanten sind deshalb bei den drei Elementen verschieden (↑ 4.2.2.5.). Die aus Spektraldaten bestimmten Rydbergkonstanten für die Elemente Wasserstoff, Deuterium und Helium verdeutlichen den Isotopieeffekt:
$R_H = (109\,677{,}575 \pm 0{,}006)\ \text{cm}^{-1}$,
$R_D = (109\,707{,}420 \pm 0{,}006)\ \text{cm}^{-1}$ und
$R_{He} = (109\,722{,}268 \pm 0{,}004)\ \text{cm}^{-1}$.
Eine Extrapolation auf unendliche Kernmasse ergibt $R_\infty = 109\,737{,}311\ \text{cm}^{-1}$. An den angegebenen Werten für die Rydbergkonstante ist erkennbar, daß man mit spektroskopischen Methoden sehr genaue Bestimmungen der Kernmasse vornehmen kann. Es gelingt mit laserspektroskopischen Methoden sogar, die Spektralverschiebungen der Elektronenterme zu beobachten, die durch die Massseänderung des Kerns Δm_k infolge einer Kernanregung ΔE gemäß der Beziehung $\Delta m_k = \frac{1}{c^2}\,\Delta E$ verursacht werden.
Nach Mehrfachionisierung konnten auch bei schwereren Atomen wasserstoffähnliche Spektren im Spektralbereich des Vakuumultraviolett oder im Röntgengebiet beobachtet werden.

4.3.3.2.3. *Röntgenspektren.* Bei Atomen höherer Ordnungszahl sind die Elektronen der Atomhülle in einzelnen Schalen angeordnet (s. a. 4.2.3.). Dabei wird die energetische Lage der von den inneren Elektronen besetzten tiefliegenden Niveaus vor allem durch die Hauptquantenzahl n bestimmt. Die energetisch niedrigste Schale ($n = 1$) wird als K-Schale, die nächst höheren ($n = 2, 3, 4, 5, \ldots$) werden als L-, M-, N-, O-, ... Schalen bezeichnet.
Die Spektren im sichtbaren und ultravioletten Spektralbereich entstehen durch Anregung äußerer Elektronen in noch höhere unbesetzte Schalen. Die Energiedifferenzen zwischen den inneren Schalen, z. B. zwischen der K- und L-Schale, sind erheblich größer und entsprechen Übergängen im Röntgengebiet. Entsprechendes Licht kann aber normalerweise nicht absorbiert oder emittiert werden, weil diese Schalen voll besetzt sind, so daß kein Elektron der K-Schale in die L-Schale oder umgekehrt übergehen kann. Röntgenspektren können deshalb nur dann beobachtet werden, wenn vorher freie Plätze in den inneren Schalen geschaffen werden.
In einer Röntgenröhre entstehen die Röntgenstrahlen durch den Aufprall stark beschleunigter Elektronen auf die Anode. Neben dem vom Anodenmaterial unabhängigen Spektrum der *Röntgenbremsstrahlung* wird diesem überlagert das sogenannte *charakteristische Röntgenspektrum* beobachtet. Die im charakteristischen Röntgenspektrum auftretenden *Röntgenemissionslinien* sind typisch für jedes Anodenmaterial. Sie entstehen auf folgende Weise: Die energiereichen auf die Anode aufprallenden Elektronen bewirken Übergänge der inneren Elektronen in höhere unbesetzte Niveaus oder in das Ionisationskontinuum. Wird z. B. ein Elektron aus der K-Schale ionisiert oder in eine hohe unbesetzte Schale gehoben, so kann die in der K-Schale entstandene Elektronenlücke durch Übergang eines Elektrons aus höheren

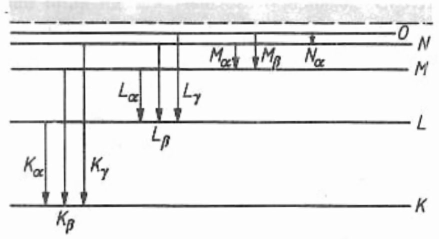


Abb. 4.3.3-2 Röntgenemissionslinien

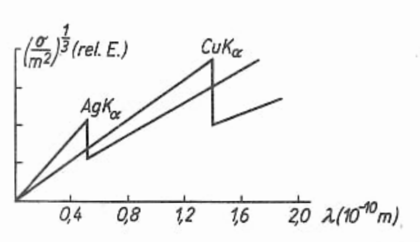


Abb. 4.3.3-3 Wellenlängenabhängigkeit des Absorptionsquerschnitts σ im Röntgengebiet

Schalen unter Emission eines Photons wieder gefüllt werden. Alle Emissionslinien, die durch Übergang aus höheren Schalen in die K-Schale entstehen, gehören zur *K-Serie*. Die einzelnen Spektrallinien werden als K_α-, K_β-, K_γ-, ... Linien bezeichnet (Abb. 4.3.3-2). Nach Emission der K_α-Linie entsteht eine Elektronenlücke in der L-Schale. Diese kann durch Übergänge aus den M-, N-Schalen gefüllt werden. Die mit diesen Übergängen aus höheren Schalen in die L-Schale verbundene Spektralserie heißt *L-Serie*.
Wie bei den Spektrallinien von Einelektronenionen (↑ 4.3.3.2.2.) im optischen Spektralbereich verschieben sich auch die Röntgenemissionslinien mit wachsender Ordnungszahl der Atome nach kürzeren Wellenlängen. MOSELEY (1913) konnte nachweisen, daß zwischen den Wellenlängen λ_{K_α} der K_α-Linien und der Ordnungszahl Z der Elemente der Zusammenhang $1/\lambda_{K_\alpha} \sim (Z-1)^2$ besteht. Ähnliche Gesetzmäßigkeiten wurden auch für andere Röntgenemissionslinien gefunden.
Während im optischen Spektralbereich die Absorptionslinien im Absorptionsspektrum bei den Wellenlängen erscheinen, wo auch die Emissionslinien im Emissionsspektrum auftre-

ten, unterscheiden sich die *Röntgenemissions- und Röntgenabsorptionsspektren* sehr stark in ihrer Struktur. Im Röntgenabsorptionsspektrum erscheinen keine scharfen Linien, sondern relativ breite Absorptionsgebiete, sogenannte *Absorptionskanten* (Abb. 4.3.3-3).

Eine Erklärung dieses Verhaltens wurde von KOSSEL gegeben. Absorbiert werden kann nur Röntgenstrahlung, deren Photonenenergie $h\nu$ so groß ist, daß Elektronen aus inneren Schalen in das Seriengrenzkontinuum bzw. in die unter der Seriengrenze liegenden unbesetzten Niveaus überführt werden können. Eine Absorption durch die Atome bei den Wellenlängen ihrer Röntgenemissionslinien ist unmöglich, weil die Schalen, in die das Elektron springen müßte, voll besetzt sind. Die langwelligen Seiten der Absorptionskanten liegen bei den Grenzwellenlängen der Röntgenemissionsserien. Der Absorptionskoeffizient steigt in der Nähe der Grenzwellenlänge steil an und fällt nach kurzen Wellenlängen hin mit λ^3 ab.

Durch den Absorptionsprozeß entsteht in den inneren Schalen eine Elektronenlücke, die durch einen Elektronenübergang aus höheren Schalen bei gleichzeitiger Röntgenemission wieder gefüllt wird. Die Absorption von Röntgenstrahlung wird also begleitet von der Emission der charakteristischen Röntgenstrahlung. Die Röntgenemissionslinien liegen auf der langwelligen Seite der Absorptionskante, d. h., es tritt eine *Röntgenfluoreszenz* auf.

Bei spektroskopischen Untersuchungen im Röntgengebiet mit hoher Auflösung ist auf der langwelligen Seite der Absorptionskante eine Strukturierung feststellbar, die auf Absorptionsübergänge in die unbesetzten optischen Niveaus unterhalb der Ionisierungsgrenze zurückzuführen ist. Die optischen Niveaus liegen nur einige Elektronenvolt unterhalb der Ionisierungsgrenze, während die Ionisierungsenergie der Elektronen innerer Bahnen bei schweren Atomen im Bereich von Kiloelektronenvolt liegt.

4.3.3.2.4. *Alkalispektren.* Die Alkaliatome, die Atome der Elemente Li, Na, K, Rb und Cs, besitzen nur ein einziges äußeres Elektron, das *Leuchtelektron.* Da dieses die chemischen Eigenschaften der Elemente bestimmt, wird es auch als *Valenzelektron* bezeichnet. Die Linienspektren im sichtbaren und ultravioletten Spektralbereich entstehen durch Anregung des Leuchtelektrons. Näherungsweise kann die Bewegung des Leuchtelektrons als eine Bewegung in einem durch die Elektronen der inneren Schalen teilweise abgeschirmten Feld des Atomkerns behandelt werden. Den Atomkern mit den abgeschlossenen Elektronenschalen bezeichnet man als *Atomrumpf.* Während beim Wasserstoffatom die Energien durch Angabe der Hauptquantenzahl eindeutig bestimmt sind, hängt die Energie des Leuchtelektrons in Alkaliatomen sowohl von der Hauptquantenzahl n als auch von der Bahndrehimpulsquantenzahl l ab (↑ 4.2.3.2.). Daher bewirkt die Auswahlregel $\Delta l = \pm 1$ eine wesentliche Einschränkung; es können nur Energieniveaus mit benachbarten l-Werten kombinieren. Dadurch wird die Zahl der möglichen Spektrallinien wesentlich verringert. In Abb. 4.3.3-4 wird

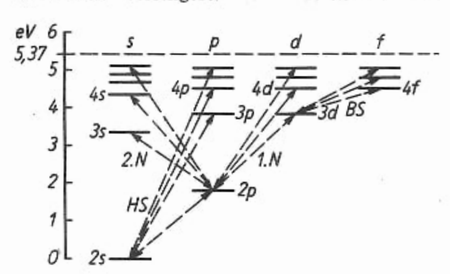


Abb. 4.3.3-4 Energieniveauschema des Lithiumatoms

dies am Beispiel des Energietermschemas des Lithiumatoms verdeutlicht. Die gestrichelt eingezeichneten Pfeile kennzeichnen die Elektronenübergänge, aus denen die vier bekannten Spektralserien der Alkaliatome resultieren. Sie tragen die Bezeichnungen Hauptserie (HS), 1. und 2. Nebenserie (1 N., 2. N.) und Bergmannserie (BS). Die Seriengesetze haben die Struktur der Rydbergformel (↑ 4.3.3.2.). Im Unterschied zum H-Atom sind die Konstanten a und b bei den Alkaliatomen ungleich Null.

In Gasentladungen hoher Stromdichten werden aber auch Emissionsspektrallinien beobachtet, die von $s \rightarrow s$- oder $s \rightarrow d$-Übergängen herrühren. Die Auswahlregel wird wegen der Einwirkung von elektrischen Feldern, die von Nachbarionen oder freien Elektronen im Plasma hervorgerufen werden, auf die angeregten Alkaliatome verletzt.

4.3.3.2.5. *Struktur der Spektren bei mehreren Valenzelektronen, Interkombinationsverbot.* Bei Betrachtung der Spektren (Tafel 67) kann man feststellen, daß mit zunehmender Zahl der Valenzelektronen die Spektrallinienzahl ansteigt. Vergleichsweise einfach sind noch die Verhältnisse im Fall des einfachsten Mehrelektronenatoms, des Heliumatoms (↑ 4.2.4.3.). Wie bei allen Atomen mit zwei Außenelektronen sind im Heliumspektrum zwei getrennte Termsysteme mit s-, p-, d-, f-, ... Termserien zu beobachten. Die beiden Termsysteme sind durch unterschiedlichen Gesamtspin, $S = 0$ bzw. $S = 1$, der beiden Außenelektronen gekennzeichnet (↑ 4.2.4.3.). Optische Übergänge zwischen beiden Termsystemen sind wegen des *Interkombinationsverbots* $\Delta S = 0$ (↑ 4.3.3.1.) verboten. Bei schweren Atomen ist das Inter-

kombinationsverbot wegen der relativ starken Spin-Bahn-Kopplung (↑ 4.2.4.4.) nicht mehr streng gültig.

Spektren mit hoher Auflösung zeigen eine *Feinstruktur der Spektrallinien.* Danach sind die Energieterme des He-Atoms mit $S = 0$ einfach, während die Terme mit $S = 1$ (außer den Zuständen mit $L = 0$) in drei energetisch dicht benachbarte Niveaus aufspalten (↑ 4.2.4.4.). Man spricht von Singulett- und Triplettserien des He-Atoms. Singulett- und Triplettserien treten bei allen Atomen mit zwei Valenzelektronen auf. Atome mit drei Außenelektronen besitzen ebenfalls zwei nicht interkombinierende Termsysteme: ein Dublettsystem ($S = 1/2$) und ein Quartettsystem ($S = 3/2$). Drei Termsysteme, nämlich Singulett- ($S = 0$), Triplett- ($S = 1$) und Quintettsysteme ($S = 2$), werden bei allen Atomen mit vier Valenzelektronen beobachtet. Typisch ist der Wechsel von gerad- und ungeradzahliger *Multiplizität* bei Erhöhung der Leuchtelektronenzahl um Eins (s. a. 4.2.4.4.).

Bei Atomen mit kleiner Kernladungszahl sind die Multiplettaufspaltungen sehr gering, so daß bei ihnen noch eine Serienstruktur im Spektrum zu erkennen ist. Mit wachsender Ordnungszahl nimmt die Termaufspaltung stark zu. Es kommt zur Überlagerung verschiedener Multipletts. Deshalb ist bei schweren Atomen eine Serienstruktur nur schwer erkennbar.

4.3.3.3. Aufspaltung der Spektrallinien. Mit genügend hochauflösenden Spektralapparaten können Aufspaltungen der Spektrallinien nachgewiesen werden, die bei Beachtung der entsprechenden Auswahlregeln unmittelbar aus den jeweiligen Aufspaltungen der atomaren Energieterme folgen. Die wesentlichsten Effekte und ihre Ursachen sind im folgenden beschrieben.

4.3.3.3.1. *Feinstruktur.* Diese wird durch die Spin-Bahn-Kopplung (↑ 4.2.4.4.) hervorgerufen und führt zu der bereits erwähnten Multiplizität der Terme (↑ 4.3.3.2.5.). Am einfachsten ist die Dublettaufspaltung bei Atomen mit nur einem Leuchtelektron. Alle Energieniveaus, außer die mit $l = 0$, spalten in zwei Niveaus auf. In Abb. 4.3.3-5 sind die Übergänge, die zur Entstehung des Na-D-Dubletts führen, eingezeichnet (s. a. Na-Spektrum in Tafel 67). Im allgemeinen sind optische Übergänge zwischen

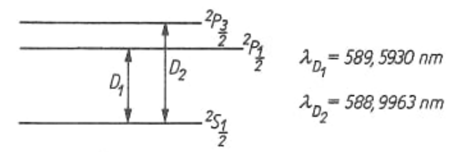


Abb. 4.3.3-5 Erklärung der Feinstruktur der Na-D-Linie

Niveaus verschiedener Multipletts nur dann erlaubt, wenn die Auswahlregel $\Delta I = 0, \pm 1$ (↑ 4.3.3.1.) erfüllt ist. Beim einfachen Beispiel des Na-D-Dubletts ergeben sich daraus allerdings keine Einschränkungen.

4.3.3.3.2. *Hyperfeinstruktur.* Die einzelnen Feinstrukturkomponenten können selbst wieder aus mehreren Linien bestehen. Diese Hyperfeinstruktur der Atomspektren beruht auf der Wechselwirkung der Elektronen mit dem Kernspin (↑ 4.2.4.8.). Genaue spektroskopische Messungen zeigen, daß die D_1-Linie des Natriums ihrerseits aus zwei Spektrallinien mit einem Wellenlängenabstand von 0,0023 nm besteht. Die D_2-Linie des Natriums spaltet ebenfalls in ein Dublett mit einem Abstand der Linien von 0,0021 nm auf. Beim Grundniveau $^2S_{1/2}$ des Natriumatoms ist der Gesamtdrehimpuls der Hülle nur durch den Spin des Leuchtelektrons gegeben. Der Kernspin des Na-23-Kerns beträgt $i = 3/2$. Das Grundniveau $^2S_{1/2}$ spaltet deshalb in die durch die beiden Atomspins $F = 3/2 - 1/2 = 1$ und $F = 3/2 + 1/2 = 2$ gekennzeichneten Unterniveaus auf (Abb. 4.3.3-6). Die Aufspaltung der angeregten Niveaus $^2P_{1/2}$ und $^2P_{3/2}$ ist wesentlich geringer als die des Grundzustands $^2S_{1/2}$. Die optischen

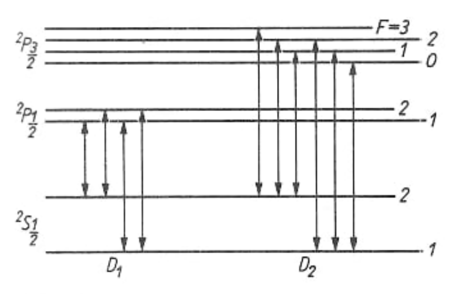


Abb. 4.3.3-6 Erklärung der Hyperfeinstruktur der Na-D-Linien

Übergänge zwischen Hyperfeinstrukturtermen werden durch die Auswahlregel $\Delta F = 0, \pm 1$ (↑ 4.3.2.1.) für die Quantenzahl F des Atomspins eingeschränkt. Die Messungen der Anzahl, der Energieabstände und der Intensitäten von verschiedenen Hyperfeinstrukturlinien gestatten die Bestimmung des Kernspins i und des magnetischen Kernmoments.

Auch *Isotopieeffekte* führen zu Hyperfeinstruktureffekten (↑ 4.2.4.9.). Es ist möglich, aus der Zahl der Hyperfeinstrukturlinien, ihren Energieabständen und Intensitäten die Anzahl und Häufigkeit der Isotope eines Elements zu bestimmen. Die Übergangsfrequenzen zwischen den Hyperfeinstrukturniveaus eines Elektronenzustands liegen im allgemeinen im Radiofrequenzbereich. Radiowellenspektrometer ge-

statten deshalb die Vermessung der Übergangsfrequenzen und Einsteinkoeffizienten für Absorptionsprozesse zwischen Hyperfeinstrukturniveaus des Atomgrundzustands.

4.3.3.3.3. *Zeemanaufspaltung.* Diese entsteht, wenn durch ein äußeres Magnetfeld die Gleichwertigkeit aller Raumrichtungen zerstört wird. Die durch die Quantenzahl M charakterisierten verschiedenen Einstellmöglichkeiten des Hüllendrehimpulses und damit auch des mit ihm verbundenen magnetischen Moments entsprechen dann $2I + 1$ unterschiedlichen Energien.

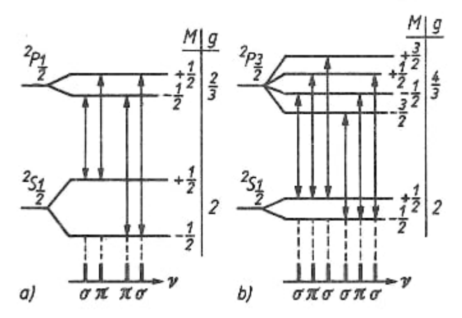


Abb. 4.3.3-7 Zeemanaufspaltung a) der Na-D_1-Linie, b) der Na-D_2-Linie; M magnetische Quantenzahl, g Landéfaktor

Die Aufspaltungen sind proportional zum Magnetfeld; bei einer magnetischen Feldstärke H von 10^4 Am^{-1} ergeben sich Energiedifferenzen von $\Delta E \approx 10^{-6}$ eV. In Abb. 4.3.3-7 sind die Zeemanaufspaltungen der beiden Natrium-D-Linien dargestellt. Optische Übergänge finden nur statt, wenn die Auswahlregel $\Delta M = 0$ oder ± 1 erfüllt ist. Der Polarisationszustand der emittierten Strahlung ist für verschiedene Übergänge unterschiedlich. Übergänge mit $\Delta M = 0$ führen zur Ausstrahlung von linear und parallel zu den Feldlinien polarisiertem Licht (π-Komponenten). Aus Emissionsübergängen mit $\Delta M = \pm 1$ resultiert rechts und links um die Feldlinien zirkular polarisierte Strahlung (σ-Komponenten). Über die Zeemanaufspaltung wird es möglich, Landéfaktoren (↑ 4.2.4.5.) zu bestimmen. Mit Hilfe der Methode der Elektronenspinresonanz (↑ 4.4.2.) ist es möglich, direkt die Absorption von Strahlung zwischen einzelnen Zeemanniveaus des Atomgrundzustands zu beobachten.

4.4. Spinresonanz

4.4.1. Allgemeine Grundlagen

4.4.1.1. Resonanz mit einem äußeren Hochfrequenzfeld. Elektronen und Nukleonen besitzen einen Spin genannten Drehimpuls (↑ 2.1.3.9.) und damit verbunden ein magnetisches Dipolmoment (↑ 2.2.2.5.), aus denen sich der Gesamtspin bzw. das Gesamtmoment der Elektronenhülle oder des Atomkerns zusammensetzt, wobei diese durch gegenseitige Kompensation auch Null ergeben können. In einem äußeren Magnetfeld $\boldsymbol{B}_0$ verhalten sich paramagnetische Teilchen, d. h. Teilchen mit einem nicht verschwindenden Gesamtspin, wie Kreisel und vollführen eine als *Larmorpräzession* bezeichnete Bewegung um die Feldrichtung (↑ 2.2.3.5.). Wie die Untersuchungen des Zeemaneffekts (↑ 4.2.4.6.) und der Stern-Gerlach-Versuch (↑ 1.1.2.5.) zeigen, richten sich die Spins im Feld aus, ihre Einstellung ist jedoch nicht beliebig möglich, sondern gequantelt (↑ 2.5.3.). Bei zuerst von Rabi durchgeführten Molekularstrahlexperimenten konnte gezeigt werden, daß es zwischen der Larmorfrequenz und einer äußeren Hochfrequenzeinstrahlung zur Resonanz und einer entsprechenden Energieübertragung auf die Spins kommen kann. Dieser Effekt läßt sich auch an stationären Proben nachweisen und hat als *Spinresonanz* eine Vielfalt von Untersuchungsmöglichkeiten eröffnet.

Im allgemeinen werden paramagnetische Teilchen untersucht, bei denen sich für die Einstellung des Spins $2I + 1$ Möglichkeiten ergeben, analog zur Zeemanaufspaltung, wobei sich die Spinvektorkomponente in Feldrichtung nur um ganze Vielfache von $\hbar$ ändern kann. Diese Einstellungen sind energetisch nicht gleichwertig; sie entsprechen den Energieniveaus, die das Spinsystem annehmen kann. Für die Übergänge zwischen den Niveaus gelten bestimmte Auswahlregeln, meist findet man $|\Delta M_I| = 1$.

Strahlt man in das System senkrecht zum Feld $\boldsymbol{B}_0$ eine Hochfrequenzwelle mit der *Larmorfrequenz* $\omega_L = \gamma B_0$ ein, so tritt Resonanz zwischen dieser und den präzedierenden Spins ein und es werden Übergänge zwischen den Energieniveaus verursacht, die Spins „klappen um". Die Energiedifferenz für einen Übergang ist $\Delta E = \hbar\omega_L \Delta M_I = \hbar\gamma B_0 \Delta M_I$. Hierbei ist das *gyromagnetische Verhältnis* γ der Quotient aus magnetischem Moment und Spin. Im thermischen Gleichgewicht sind die niederenergetischen Niveaus des Spinsystems stärker besetzt als die höheren, daher werden hauptsächlich Übergänge in die höheren Niveaus induziert, es kommt zur Energieaufnahme aus dem Hochfrequenzfeld. Auf dieser Energieabsorption beruhen nahezu alle Meßverfahren der Spinresonanz.

Berücksichtigt man den Zusammenhang zwischen magnetischem Moment und Spin $\boldsymbol{m} = g\mu_x \boldsymbol{I}/\hbar$, und setzt man den daraus folgenden Wert für das gyromagnetische Verhältnis in die Energieformel ein, so erhält man ΔE

$= g\mu_x B_0 \Delta M_I$. Dabei wird g Aufspaltungs- oder g-Faktor genannt. Er hängt von der Struktur der Substanz ab und ist ein Maß für die Größe der Aufspaltung der Energieniveaus. Für die Elektronenhülle freier Atome ist g gleich dem Landéfaktor (↑ 4.2.4.5.). Wechselwirkungen zwischen Nachbaratomen führen zu Abweichungen vom Landéfaktor im Bereich von etwa 10^{-3}; der genaue Wert wird experimentell bestimmt. Für Atomkerne variiert der Kern-g-Faktor g_I je nach Kernart zwischen $-1{,}29$ und $+2{,}79$. Für μ_x ist bei Elektronen das Bohrsche Magneton μ_B (↑ 8.7.1.2.) zu setzen, bei Atomkernen verwendet man wegen der um den Faktor 1837 bis 45000 größeren Masse und dem um den gleichen Faktor geringeren magnetischen Moment das Kernmagneton $\mu_k = \mu_B/1837$.

Mit steigender Frequenz, d. h. mit anwachsendem Magnetfeld, erhöht sich die Energiedifferenz der Spinniveaus und damit die Nachweisempfindlichkeit der Resonanzabsorption. Die Spinresonanz konnte daher erst nach der Entwicklung einer leistungsfähigen Hochfrequenz- bzw. Mikrowellentechnik untersucht werden. 1945 wurde die erste Elektronenresonanzabsorption von SAWOJSKY im UKW-Gebiet mit relativ geringem Magnetfeld gemessen. Der Nachweis der Kernresonanz gelang erstmals 1946 BLOCH. Gegenwärtig liegen die benutzten Magnetfelder bei 1 T, die zugehörigen Frequenzen für die Kernresonanz im UKW-Bereich bei etwa $8 \cdot 10^7$ Hz, für Elektronenresonanz im Mikrowellengebiet um 10^{10} Hz. Man spricht daher allgemein von *Hochfrequenzspektroskopie*, und speziell bei der Elektronenresonanz von *Mikrowellenspektroskopie*.

Eng verwandt mit der paramagnetischen Resonanz sind die *akustische Resonanz*, bei der an Stelle des elektromagnetischen Hochfrequenzfelds Ultraschall eingestrahlt wird, und die *Quadrupolresonanz*, die die Orientierung von Atomkernen mit einem Quadrupolmoment in einem inhomogenen elektrischen Feld ausnutzt.

4.4.1.2. Relaxationsprozesse

4.4.1.2.1. *Relaxationszeiten.* Das in 4.4.1.1. skizzierte Meßprinzip erweist sich in der Praxis als unzulänglich. Der Nachweis der Absorption ist im stationären Betrieb schwierig, denn nach kurzer Zeit tritt Sättigung ein, d. h. alle Spinniveaus sind gleichmäßig besetzt, alle Übergänge gleichwahrscheinlich, so daß sich Absorption und Emission die Waage halten und in der Bilanz keine Energie aus dem Hochfrequenzfeld entnommen wird. Dies tritt besonders bei zu großer Amplitude dieses Felds auf. Man mißt daher entweder mit kleiner Hochfrequenzamplitude und muß erhebliche Empfindlichkeitseinbußen in Kauf nehmen oder überstreicht die Resonanz periodisch, indem man das Magnetfeld oder die Hochfrequenz moduliert und zwischen zwei Resonanzdurchgängen das Spinsystem in den Ausgangszustand zurückkehren läßt.

Wird das Spinsystem durch die Energieabsorption aus dem thermischen Gleichgewicht gebracht, versucht es durch Ausgleichsvorgänge, sogenannte *Relaxation*, zu diesem Gleichgewicht zurückzukehren. Zuerst erfolgt das innerhalb des Spinsystems mit der Zeitkonstanten T_2, der *Spin-Spin-Relaxationszeit*. Bei dieser Wechselwirkung ändert sich der Energieinhalt des Spinsystems nicht. Wesentlich langsamer erfolgt der Energieaustausch zwischen dem Spinsystem und dem Gitter (hier allgemein als Umgebung des Spinsystems verstanden, auch wenn es sich nicht um einen Festkörper handelt) mit der *Spin-Gitter-Relaxationszeit* T_1 und damit die Rückkehr zum thermischen Gleichgewicht innerhalb der gesamten Probe. Die für die zeitliche Änderung der Spinkomponente in Richtung des äußeren Felds maßgebende *longitudinale Relaxationszeit* T_1 stimmt mit der Spin-Gitter-Relaxationszeit überein. Sie ist für das Sättigungsverhalten verantwortlich, große T_1 bedeuten leichte Sättigung der Resonanz. Für die zum Feld $\boldsymbol{B}_0$ senkrechte Komponente erhält man die *transversale Relaxationszeit* T_2^*, die durch die Überlagerung der Einflüsse von Spin-Spin- und Spin-Gitter-Wechselwirkungen sowie von Inhomogenitäten des Felds entsteht. Für sie gilt daher $T_2^* \leqq T_2$. T_2^* bestimmt im wesentlichen die Breite der Resonanzlinien.

4.4.1.2.2. *Impulssättigungsmethode.* Zur direkten Messung der Relaxationsparameter werden Hochfrequenzimpulsmethoden angewandt. Das Spinsystem wird dabei aus dem thermischen Gleichgewicht mit seiner Umgebung gebracht und anschließend beobachtet, wie es in den Gleichgewichtszustand zurückkehrt. Bei der *Impulssättigungsmethode* läßt man auf die Probe im konstanten Magnetfeld B_0 einen Hochfrequenzimpuls B_1 einwirken, dessen Frequenz der Resonanz entspricht und dessen

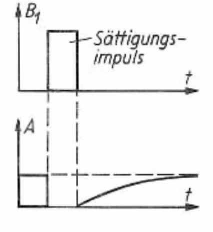


Abb. 4.4.1-1 Wiederanwachsen der Signalamplitude A nach einem Sättigungsimpuls; B_1 Hochfrequenzimpuls

Intensität so groß ist, daß er das Spinsystem sättigt. Ein kontinuierlich eingestrahltes Feld der gleichen Frequenz kleiner Amplitude, ein sogenanntes Monitorfeld, zeigt erst nach Ab-

klingen der Sättigung ein Wiederanwachsen des Absorptionssignals (Abb. 4.4.1-1) mit der Zeitkonstanten T_1. Das Monitorfeld selbst darf keinerlei Sättigung bewirken. Die Empfindlichkeit kann dadurch gesteigert werden, daß man auch das Monitorfeld nur für eine kurze Beobachtungszeit impulsförmig mit größerer Amplitude einstrahlt.

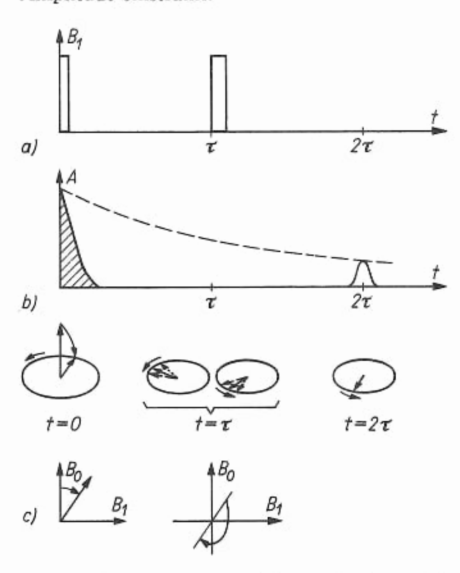


Abb. 4.4.1-2 Spinecho; a) Impulsschema der $\boldsymbol{B}_1$-Impulse zur Messung von T_2, b) Signale: freie Induktion *schraffiert*, Spinecho und Zeitabhängigkeit der Echoamplitude *gestrichelt*, c) perspektivische Darstellung der Spinbewegung mit Angabe der Feldrichtungen

4.4.1.2.3. *Spinechomethoden.* Dabei wirken auf die im konstanten Feld befindliche Probe mehrere Hochfrequenzimpulse mit definierter Länge und Feldstärke ein. Man verwendet Kombinationen aus sogenannten $\pi/2$- und π-Impulsen, die eine Drehung der Spins gegenüber der Feldrichtung um 90° bzw. 180°, d. h. im Bogenmaß um $\pi/2$ bzw. π bewirken. Abb. 4.4.1-2 veranschaulicht schematisch die Methode zur T_2^*-Bestimmung. Auf die im konstanten Feld ausgerichteten Spins wirkt zur Zeit $t = 0$ ein $\pi/2$-Impuls und klappt sie um 90° in die zu $\boldsymbol{B}_0$ senkrechte Ebene, in der sie um die $\boldsymbol{B}_0$-Richtung präzedieren und in einer zu $\boldsymbol{B}_0$ senkrechten Spule ein als freie Induktion bezeichnetes Signal induzieren. Durch Relaxationseinflüsse und Feldinhomogenitäten ist die Präzessionsgeschwindigkeit der einzelnen Spins unterschiedlich, die Spinrichtungen laufen auseinander und die freie Induktion klingt ab. Nach einer Zeit $t = \tau$ wird ein π-Impuls eingestrahlt, der alle Spins in der Präzessionsebene um B_1 als Achse um 180° klappt. Damit eilen die schneller präzedierenden Spins den langsameren nach. Nach Verlauf der gleichen Zeit τ, d. h. bei $t = 2\tau$, befinden sich die Spins wieder in einheitlicher Richtung und induzieren das sogenannte *Spinecho*. Durch Veränderung des Impulsabstands τ erhält man die Echoamplitude in Abhängigkeit von der Zeit und kann daraus direkt T_2^* ableiten, die Zeitkonstante für den Abbau der transversalen Induktion. Mit anderen Impulsfolgen läßt sich auch T_1 bestimmen, z. B. mit $\pi - \pi/2 - \pi$. Hier werden zuerst die Spins um 180° umgeklappt, d. h. entgegengesetzt zur Feldrichtung, also in das höhere Energieniveau gebracht, danach erfolgt Rückkehr in den Ausgangszustand mit der Zeitkonstanten T_1. Mit dem folgenden Spinecho, den $\pi/2$-π-Impulsen, wird diese Rückkehr in Abhängigkeit von der Zeit untersucht. Man erhält zuerst negative Echos, d. h. Emissionssignale, dann positive. Aus dem Nulldurchgang läßt sich T_1 sehr genau ermitteln.

4.4.2. Paramagnetische Elektronenresonanz

4.4.2.1. Niveauaufspaltungen. Die paramagnetische Elektronenresonanz, auch als *Elektronenspinresonanz* bezeichnet, meist mit EPR [*electron paramagnetic resonance* engl.] abgekürzt, ist das Teilgebiet der Spinresonanz, das die selektive Absorption elektromagnetischer Strahlung infolge der Umorientierung durch Elektronen bedingter magnetischer Momente untersucht. Solche Prozesse treten in paramagnetischen Substanzen oder an paramagnetischen Zentren auf, wenn ein Magnetfeld angelegt wird. Zu den magnetischen Momenten tragen sowohl der eigentliche Spin als auch der Bahndrehimpuls der Elektronen bei. Diese setzen sich zum Elektronengesamtspin J zusammen. Die EPR ist ein Spezialfall der Zeemanspektroskopie (↑ 4.3.3.3.3.), bei der die Übergänge zwischen Zeemanniveaus direkt gemessen werden und nicht nur als kleine Aufspaltungen bei optischen Untersuchungen auftreten.

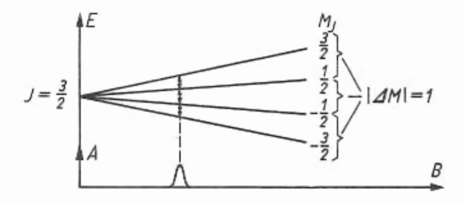


Abb. 4.4.2-1 Aufspaltung des Spinniveaus $J = 3/2$ im äußeren Feld B; A Amplitude des Resonanzsignals, E Energie

Ohne äußeres Feld sind die Energieniveaus der Elektronen freier Atome und paramagnetischer Stoffe entartet, d. h. die verschiedenen Spineinstellungen haben die gleiche Energie. Bei Anlegen eines Felds spalten die Niveaus gleichmäßig auf, wegen der Auswahlregel $|\Delta M_J| = 1$ ergibt sich für alle Übergänge die gleiche Linie (Abb. 4.4.2-1).

In den meisten Fällen tritt jedoch nicht nur *eine* EPR-Linie auf, sondern mehrere, oft in Gruppen. Eine sogenannte *Feinstruktur* findet man, wenn der Gesamtspin der beteiligten Elektronen $J > 1/2$ ist. Hierbei wird durch innere Felder der Substanz die bei isolierten Atomen vorliegende Entartung teilweise aufgehoben, so daß auch ohne äußeres Feld bereits unterschiedliche Niveaus entstehen, man spricht von *Null-* oder *Kristallfeldaufspaltung*. Die gleiche Energiedifferenz $\Delta E = h\nu$ für einen Übergang $|\Delta M_J| = 1$ und damit die gleiche Resonanzfrequenz erhält man für verschiedene Feldwerte, es erscheinen beim Verändern des Felds nacheinander mehrere Linien (Abb. 4.4.2-2). Unter der Voraussetzung $|\Delta M_J| = 1$ ergeben sich aus den $2J + 1$ Zuständen $2J$ Feinstrukturlinien. Durch Messung bei verschiedenen Frequenzen kann der Verlauf dieser Aufspaltung in Abhängigkeit von der Feldstärke bestimmt werden. Häufig beobachtet man eine noch feinere Aufspaltung der Resonanzlinien, die *Hyperfeinstruktur* (↑ 4.2.4.8.). Sie tritt immer dann auf, wenn die Atomkerne einen von 0 verschiedenen Kernspin I besitzen, der mit den Elektronenspins J koppelt. Bei den in der Praxis verwendeten starken Magnetfeldern stellen sich I und J getrennt im Feld ein, analog wie Elektronenspin und -bahndrehimpuls beim Paschen-Back-Effekt (↑ 4.2.4.6.), so daß jeder Feinstrukturzustand in $2I + 1$ Hyperfeinstrukturzustände aufgespalten ist (Abb. 4.4.2-3). In derartig starken Magnetfeldern gilt die Auswahlregel $\Delta M_I = 0$, d. h., die Kernorientierung bleibt erhalten. Damit wird die Zahl der Hyperfeinstrukturlinien ebenfalls $2I + 1$. Insgesamt treten durch Fein- und Hyperfeinstruktur $2J(2I + 1)$ Linien auf, z. B. 12 Linien für $I = 3/2$, $J = 3/2$. Meist stehen die beteiligten Elektronen nicht nur mit einer, sondern mit mehreren Kernarten mit unterschiedlichem Kernspin in Wechselwirkung, so daß weitere Aufspaltungen entstehen.

4.4.2.2. Linienverbreiterungen. Die natürliche Linienbreite der EPR ist der mittleren Lebensdauer der Besetzung des Spinzustands umgekehrt proportional, das bedeutet, je länger die Lebensdauer ist, um so schmaler ist die zugehörige Linie. Die für freie Atome berechneten natürlichen Linienbreiten sind mit größenordnungsmäßig 10^{-4} Hz weitaus geringer als die experimentell beobachteten. Diese Linienverbreiterungen können verschiedene

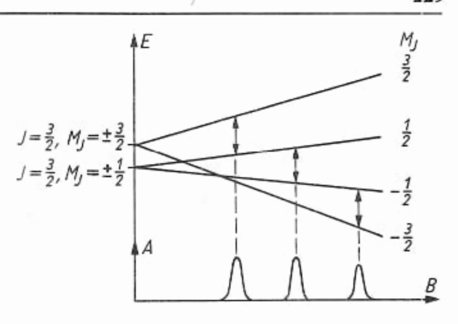


Abb. 4.4.2-2 Feinstruktur infolge Nullfeldaufspaltung

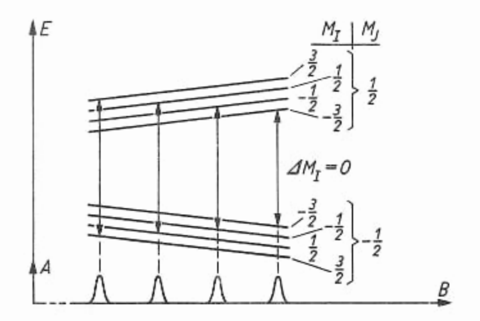


Abb. 4.4.2-3 Hyperfeinstruktur bei Kernspin $I = 3/2$, Elektronenspin $J = 1/2$

Ursachen haben. Die *Dipol-Dipol-Wechselwirkung*, auch *Spin-Spin-Wechselwirkung* genannt, zwischen paramagnetischen Nachbaratomen verursacht Lokalfelder, die dem äußeren Feld überlagert sind und dieses ortsabhängig ändern. Dieser Einfluß kann durch „Spinverdünnung" mit diamagnetischen Teilchen verringert werden. Die *Spin-Gitter-Wechselwirkung* wird durch die Spin-Gitter-Relaxationszeit T_1 (↑ 4.4.1.2.) bestimmt. Kleine T_1 ergeben breite, niedrige Linien; große führen zur Liniensättigung. T_1 ist in hohem Maße durch die Umgebung bestimmt und kann schon durch geringe Beimischungen oder Verunreinigungen der Substanz stark geändert werden. Weitere Ursachen für Linienverbreiterungen sind nichtaufgelöste Hyperfeinstrukturen, Sättigung der Linien, Inhomogenitäten des angelegten Magnetfelds und des Kristallaufbaus, thermische Gitterschwingungen und Kristallanisotropie.

4.4.2.3. Anwendungen der EPR

4.4.2.3.1. *Nachweis.* Er erfolgt über die Bestimmung der Absorption der eingestrahlten Mikrowellenenergie. Meist verwendet man eine *Mikrowellenbrücke*, bei der die Strahlung in zwei gleiche Mikrowellenresonatoren einge-

speist wird, die so abgestimmt sind, daß das Differenzsignal Null ist. Einer der Resonatoren enthält die Probensubstanz, und bei Resonanzabsorption ändert sich dessen Güte, die Brücke wird verstimmt, und am Ausgang erhält der sogenannte Spindetektor ein Signal. Die Brükkenschaltung hat den Vorteil, daß auch bei großer Hochfrequenzamplitude kleine Absorptionssignale nachgewiesen werden können.

4.4.2.3.2. *Untersuchung freier Radikale.* In der flüssigen Phase erhält man neben spektroskopischen Daten Aussagen über deren chemische Reaktivität, die sich im mehr oder weniger schnellen Abbau der Radikalkonzentration äußert. Bei stationärer Registrierung des EPR-Spektrums mißt man diese Konzentration im dynamischen Gleichgewicht zwischen Radikalerzeugung und -rekombination und kann bei Kenntnis der chemisch oder strahleninduzierten Radikalbildungsrate relative Reaktivitäten bestimmen. Für kurzlebige Radikale, wie Zwischenprodukte in chemischen Reaktionsabläufen, ist jedoch die *kinetische* oder *zeitaufgelöste EPR* weitaus besser geeignet. Hierbei wird die Probe impulsförmig zur Bildung der interessierenden Radikale angeregt, z. B. durch Elektronenbeschuß, Laserbestrahlung, und die Konzentration der Radikale im Nichtgleichgewicht über EPR nachgewiesen. Notwendig ist dafür eine Modifikation der EPR mit einer Zeitauflösung von einigen Mikrosekunden. Zur Untersuchung von Reaktionsabläufen, Relaxationszeiten usw. verwendet man häufig Spinechomethoden (↑ 4.4.1.2.).

4.4.2.3.3. *Ferromagnetische Elektronenresonanz.* Auch in bestimmten ferromagnetischen Stoffen läßt sich Spinresonanz nachweisen. Ferromagnetika besitzen ein starkes inneres Feld, das bis zu 10^3 T betragen kann und damit weit stärker ist als das äußere, so daß sich die Resonanzbedingungen gegenüber der normalen EPR verschieben. Infolge der starken Kopplung der Spins präzedieren diese in der gesamten Probe gemeinsam und gleichphasig, man erhält daher ein großes makroskopisches magnetisches Moment und große Signalintensitäten. Die Untersuchungen dieser *ferromagnetischen Elektronenresonanz* gestatten Aussagen über die Struktur und die inneren Felder von Ferromagnetika, über die in ihnen vorliegenden Wechselwirkungen und aus der Linienbreite Aussagen über den Dämpfungsmechanismus.

Die EPR wird meist zur Untersuchung freier Radikale und paramagnetischer Zentren benutzt. Beispiele sind die Untersuchung von Kristallstrukturen, Bestimmung der Valenzzustände eingebauter Ionen, Untersuchung kovalenter Bindungen, besonders in der chemischen und biologischen Forschung. Bei bekannter Resonanzfrequenz der Probe lassen sich auch Magnetfelder mit hoher Genauigkeit ausmessen.

4.4.3. Paramagnetische Kernresonanz

4.4.3.1. Grundlagen. Ein weiteres Hauptgebiet der Spinresonanz ist die paramagnetische Kernresonanz, die auch Kernspinresonanz genannt wird und meist mit NMR [*nuclear magnetic resonance* engl., kernmagnetische Resonanz] abgekürzt. Hier wird die Einstellung der Kernspins im äußeren Magnetfeld untersucht. Die Relaxationswechselwirkungen sind für die Kernspins erheblich geringer als bei EPR, so daß sich erheblich größere Relaxationszeiten ergeben. Die longitudinale Relaxationszeit T_1 (↑ 4.4.1.2.) liegt bei reinen Flüssigkeiten im Bereich von Sekunden, in Kristallen bei einigen Tagen. Durch paramagnetische Beimischungen kann T_1 erheblich verkürzt und die Sättigung verringert werden.

In Flüssigkeiten ist die transversale Relaxationszeit $T_2^* \approx T_1$ und damit im Sekundenbereich, für Festkörper gilt $T_2^* \ll T_1$ mit T_2^* in der Größenordnung 10^{-5} s. In Festkörpern findet man daher im allgemeinen relativ breite Linien, die dadurch verursacht werden, daß sich die Felder der statistisch orientierten Kerndipole im Kristallgitter überlagern und ein räumlich ungleichmäßiges lokales Feld bilden, in dem die einzelnen Kerne unterschiedliche Resonanzfrequenzen aufweisen. Untersuchungen dieser Art bezeichnet man als *Breitlinienspektroskopie*.

In Metallen und Halbleitern beobachtet man eine Verschiebung der Resonanzfrequenz gegenüber freien Atomen infolge der Wechselwirkung der Kernspins mit den magnetischen Momenten der Leitungselektronen, man spricht von Knight-Shift [*shift* engl., Verschiebung].

4.4.3.2. Hochauflösende Kernresonanz. In Flüssigkeiten werden die inneren magnetischen Felder am Ort der Kernspins durch die schnelle Umorientierung und Diffusion der Moleküle weitestgehend ausgemittelt. Die beobachteten Linien sind daher sehr schmal, mit etwa 10^{-7} bis 10^{-6} T um einige Größenordnungen geringer als bei Festkörpern. Die zugehörigen Meßverfahren bezeichnet man als *hochauflösende Kernresonanz*. Dabei sind zwei Phänomene charakteristisch: die chemische Verschiebung und die *indirekte Spin-Spin-Kopplung*. Die *chemische Verschiebung* resultiert daraus, daß das Magnetfeld am Kernort durch die den Kern umgebenden Elektronen teilweise kompensiert wird und sich auf $B' = B_0(1 - \sigma)$ verringert. σ ist die Abschirmkonstante. Für gleiche Kerne in unterschiedlicher chemischer Umgebung wird das Feld verschieden stark abgeschirmt, so daß sich unterschiedliche Resonanzfrequenzen ergeben, die zu einer Linienaufspaltung nach Abb. 4.4.3-1a führen. Die Lage der Linien wird immer auf eine Standardsubstanzlinie bezogen; Maß für die

chemische Verschiebung ist $\delta = (\nu - \nu_s)/\nu_s$, wobei ν_s die Resonanzfrequenz der Standardsubstanz ist. δ liegt für Protonenresonanzen bei 10^{-6} bis 10^{-5}. Die durch die chemische Verschiebung entstehenden Linien spalten oft

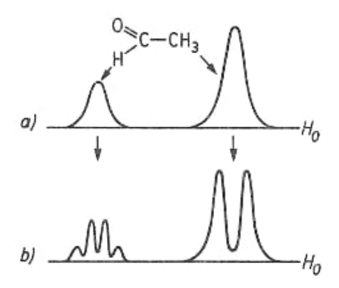


Abb. 4.4.3-1 Protonenresonanzspektrum von Acetaldehyd (CH_3CHO); a) die beiden infolge der chemischen Verschiebung auftretenden Protonenresonanzlinien, b) Spektrum bei höherer Auflösung mit Multiplettstruktur infolge indirekter Spin-Spin-Kopplung

weiter auf, sie zeigen nach Abb. 4.4.3-1b Multiplettstruktur. Diese resultiert aus einer Kopplung zwischen zwei Kernspins, die aber nicht wie im Festkörper direkt erfolgt, sondern indirekt durch Vermittlung der Valenzelektronen. Durch Austauschkräfte kommt es über eine oder mehrere Bindungen hinweg zu dieser Kopplung, die naturgemäß mit der Anzahl der beteiligten Bindungen, d. h. mit dem Abstand der Spins sinkt, so daß man Aussagen über die Größe des Moleküls und die Atomanordnung erhält. Die Multiplettaufspaltung ist von der Größe des äußeren Felds unabhängig und kann dadurch von der chemischen Verschiebung unterschieden werden.

Die hochauflösende Kernresonanz konnte auch auf Festkörper ausgedehnt werden. Um die direkte Dipol-Dipol-Kopplung zu eliminieren, wurden bisher drei Methoden entwickelt: 1. Eine schnelle Rotation der Probe im äußeren Magnetfeld unter dem sogenannten magischen Winkel 54°44′ wirkt praktisch wie die thermische Umorientierung und Diffusion in einer Flüssigkeit. 2. Man untersucht mit seltenen Kernen, z. B. C-13 in einer C-12 Probe, deren gegenseitiger Abstand so groß ist, daß die direkte Spin-Spin-Kopplung verschwindet. Die starke Empfindlichkeitseinbuße läßt sich durch Doppelresonanz (↑ 4.4.5.) ausgleichen. 3. Impulsartige Einstrahlung des Hochfrequenzfelds nach einem geeigneten Programm variiert die Spineinstellung zeitlich derart, daß sich die Dipol-Dipol-Wechselwirkung ausmittelt und nur der interessierende Anteil bleibt. Mit dieser Methode lassen sich auch die häufigen Kerne untersuchen. Die in der hochauflösenden Kernresonanz hauptsächlich verwendeten Kerne sind Protonen, seltener C-13 und F-19.

4.4.3.3. Experimenteller Nachweis und Anwendung der NMR. Der experimentelle Nachweis ist dem der EPR ähnlich, da die Untersuchungen aber im UKW-Bereich erfolgen, können die konventionellen Bauelemente der Hochfrequenztechnik, wie Spulen oder Kondensatoren verwendet werden. Mit der Methode der *Kerninduktion* konnte Bloch erstmals die Kernresonanz nachweisen. Er koppelte in die im Magnetfeld $\boldsymbol{B}_0$ befindliche Probe durch eine Spule senkrecht zu den Feldlinien ein Hochfrequenzfeld $\boldsymbol{B}_1$ ein und erhielt in einer zweiten, zu $\boldsymbol{B}_0$ und $\boldsymbol{B}_1$ senkrechten Spule im Falle der Resonanz ein Signal. Bei den zur Zeit üblichen Geräten bestimmt man die Resonanzabsorption. Die Probe befindet sich dabei im Inneren einer Spule, deren Achse senkrecht zu den Feldlinien von $\boldsymbol{B}_0$ steht. Bei Resonanz führt die Energieabsorption zu einer Bedämpfung der Spule, und ein an diese angeschlossener Oszillator verringert dann seine Schwingamplitude. Neben den stationären Methoden benutzt man in großem Maße Spinechoverfahren (↑ 4.4.1.2.). Die NMR stellt hohe Anforderungen an die Konstanz und Homogenität des Magnetfelds, bei hochauflösenden Verfahren muß $\Delta B/B < 10^{-7}$ sein. Modernste Geräte benutzen Frequenzen von über 200 MHz, die zugehörigen Felder von etwa 5 T lassen sich nur mit supraleitenden Magneten erzeugen.

Die NMR wird im allgemeinen zur Strukturanalyse verwendet, die chemische Verschiebung z. B. zur Untersuchung der räumlichen Lage der Kerne im Molekül, von Wasserstoffbrücken und von charakteristischen Gruppen im Molekül. Aus den Linienintensitäten gewinnt man Aussagen über Gemischzusammensetzungen, führt also Elementanalysen durch, aus den Linienbreiten ergeben sich Informationen über Geschwindigkeiten von Austauschprozessen. Die Relaxationszeiten liefern Aussagen über die Bewegung des Moleküls und seiner Teile, sowie über die Kinetik schneller reversibler Reaktionen. Außerdem erlauben sie Konfigurationsbestimmungen. Man erhält weiterhin die Größen der Kernspins und ist in der Lage, Magnetfelder mit höchster Genauigkeit zu messen. Die NMR wird u. a. zur Untersuchung von Biomembranen, Stoffwechselprozessen und als medizinische Diagnosemethode eingesetzt, z.. B ist das NMR-Spektrum einer Krebszelle gegenüber dem einer normalen Zelle verändert.

4.4.4. Akustische Spinresonanz

Die akustische Spinresonanz ist das Analogon der paramagnetischen Spinresonanz, man findet

bei Einstrahlung von Ultraschall ebenfalls die Anregung von Übergängen zwischen Energieniveaus, die durch die unterschiedlichen Stellungen von Spins in konstanten magnetischen und elektrischen Feldern hervorgerufen werden. Durch die Einstrahlung von Ultraschall werden die in einer Probe vorhandenen elektrischen Ladungen und magnetischen Dipole ähnlich wie durch elektromagnetische Wechselfelder zu Schwingungen angeregt. Je nach dem Frequenzbereich kann man akustische Kernresonanz (1 bis 100 MHz) und Elektronenresonanz (10^3 bis 10^4 MHz) anregen. Trotz der wesentlich geringeren Empfindlichkeit wird hauptsächlich die akustische Kernresonanz untersucht, da sich dieser Frequenzbereich weitaus einfacher technisch realisieren läßt. Die Experimente werden meist an Einkristallen durchgeführt, deren Stirnflächen optisch eben, parallel und senkrecht zur Schallrichtung sein müssen. Ein solcher Kristall wirkt als akustischer Resonator der Länge l, in dem sich stehende Wellen der Wellenlänge λ anregen lassen, für die $l = n\lambda/2$ gilt, wobei n eine ganze Zahl ist. Durch Änderung des äußeren Magnetfelds erzielt man Resonanz zwischen den mechanischen Schwingungen des Kristalls und der Spinpräzession. Der durch die Resonanz absorbierte Schallanteil ist mit etwa 10^{-5} so gering, daß der akustische Nachweis kaum möglich ist, man verwendet daher die empfindlicheren elektrischen Methoden, z. B. ist die Probe Teil eines elektromagnetischen Wandlers, der in einem elektrischen Schwingkreis eingebaut ist und bei Resonanz durch die Schallabsorption den Schwingkreis bedämpft. Meist weist man jedoch indirekt über EPR oder NMR nach, indem man die Größe der elektrischen Resonanzabsorption mit und ohne Schallfeld bestimmt. Es wurden auch Spinechomethoden (↑ 4.4.1.2.) erfolgreich eingesetzt. Trotz der größeren technischen Schwierigkeiten wird die akustische Resonanz neben EPR und NMR benutzt, da sie auch Untersuchungen von Übergängen mit $|\Delta M| = 2$ gestattet, die zusätzliche Informationen über das Spinsystem liefern. Außerdem lassen sich mit der akustischen Resonanz elektrisch leitende Proben weitaus besser untersuchen, weil der bei Hochfrequenz störende Einfluß des Skineffekts wegfällt.

4.4.5. Doppelresonanz und verwandte Effekte

Bei den Messungen der Spinresonanz werden häufig Verfahren angewendet, bei denen neben dem Hochfrequenzfeld B_1 zusätzlich Energie in das Spinsystem eingebracht wird, meist ebenfalls in Form eines Hochfrequenzfelds B_2, das zusätzlich Resonanzerscheinungen in anderen Teilen des Spinsystems hervorruft. Hierbei hat sich eine große Zahl von speziellen Techniken herausgebildet.

4.4.5.1. Kern-Kern-Doppelresonanz. Die Kern-Kern-Doppelresonanz, meist mit INDOR [*internuclear double resonance* engl., Doppelresonanz zwischen Kernen] abgekürzt, untersucht im Spinsystem zwei Kernsorten unterschiedlicher Resonanzfrequenzen, entweder gleiche Kerne, die an verschiedenen Orten eines Moleküls infolge chemischer Verschiebung (↑ 4.4.3.) abweichende Resonanzfrequenzen aufweisen, oder unterschiedliche Kerne, z. B. Protonen und C-13-Kerne. Wählt man die Amplituden der mit den beiden Resonanzfrequenzen eingestrahlten HF-Felder B_1 und B_2 klein, ergeben sich durch die Spin-Spin-Kopplung nur Änderungen in den relativen Intensitäten von Multiplettlinien. Wichtiger ist der Fall, daß B_2 so groß ist, daß die zugehörige Resonanz gesättigt wird. Dann werden die durch die Spin-Spin-Kopplung bedingten Übergänge zwischen den Energieniveaus der beiden Kernsorten für Absorption und Emission gleichwahrscheinlich, die Spin-Spin-Kopplung wirkt nach außen wie nicht vorhanden, so daß man in komplizierten Spektren bestimmte Linien gezielt eliminieren kann.

4.4.5.2. Elektron-Kern-Doppelresonanz. Bei dieser als ENDOR-Technik [*electron nucleus double resonance* engl.] bezeichneten Methode wird ein Spinsystem aus Kernspin I und Elektronenspin J, also mit Hyperfeinstruktur (↑ 4.4.2.), untersucht, wobei die Elektronenresonanz beobachtet wird. Abb. 4.4.5-1 zeigt das Beispiel eines Spinsystems mit $I = J = 1/2$. Die Niveaus entsprechen den angegebenen Spinprojektionen und sind im thermischen Gleichgewicht entsprechend der Spintemperatur besetzt. Durch ein starkes, sättigendes Mikrowellenfeld mit der Frequenz ω_e, die zur Elektronenresonanz gehört, gleicht man die Besetzungszahlen der äußeren Niveaus nahezu aus und erzeugt eine Kernpolarisation (↑ 4.4.5.4.). Ein zusätzliches Kernresonanzfeld mit der Frequenz ω_n hebt die Polarisation wieder auf und entsättigt das Elektronenspinsignal, wodurch man eine Erhöhung der Absorption erhält. Durch die Elektronen- bzw. Kernrelaxation

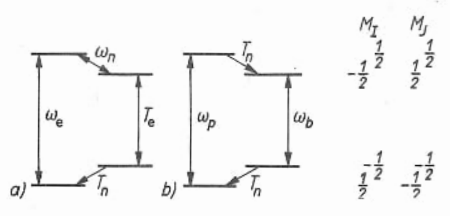


Abb. 4.4.5-1 Doppelresonanzen; Niveauschema und Übergänge a) beim ENDOR-, b) beim ELDOR-Verfahren

(↑ 4.4.1.2.) kehrt das System mit den Relaxationszeiten T_e bzw. T_n in den Ausgangszustand zurück. Der Vorteil der ENDOR-Technik liegt in der gegenüber EPR wesentlich geringeren Linienbreite, die theoretisch um den Faktor ω_e/ω_n abnimmt, experimentell wurden Verringerungen auf 10^{-2} bis 10^{-4} beobachtet. ENDOR ist oft das einzige Verfahren zur Untersuchung kleinster Hyperfeinstrukturwechselwirkungen.

4.4.5.3. Elektronendoppelresonanz. Bei diesem als ELDOR [*electron double resonance* engl.] bezeichneten Verfahren wird das gleiche Spinsystem wie bei der ENDOR-Technik untersucht, es werden jedoch zwei verschiedene Elektronenresonanzfrequenzen benutzt. Abb. 4.4.5-1 zeigt das Prinzip: Der Übergang ω_p (Pumpfrequenz) wird gesättigt, die Beobachtungsfrequenz ω_b mit niedriger Intensität untersucht. Durch die Wechselwirkung mit den Kernspins wird mit der Kernrelaxationszeit T_n Energie von einer Elektronenresonanz auf die andere übertragen und damit das Signal bei ω_b erhöht.

Das ELDOR-Verfahren eignet sich besonders zur Untersuchung dynamischer Prozesse, da die Abhängigkeit der Intensität vom Relaxationsverhalten besonders einfach ist, weiterhin zur Verbesserung der Auflösung bei großen Hyperfeinstrukturaufspaltungen und starken Austauschprozessen.

4.4.5.4. Overhausereffekt. Bei diesem nach dem Physiker OVERHAUSER benannten Effekt betrachtet man ebenfalls ein System von gekoppelten Kernspins I und Elektronenspins J, beobachtet jedoch bei einem Feld der Intensität B_1 die Kernresonanz ω_n. Erzeugt man mit B_2 zusätzlich Elektronenresonanz ω_e, steigt bei kleinem B_2-Feld infolge Spin-Spin-Wechselwirkung zwischen Kernen und Elektronen die Intensität des Kernresonanzsignals an. Bei Sättigung der Elektronenresonanz erhält man durch diese Kopplung eine Besetzungsinversion (↑ 4.1.4.1.) der Kernniveaus. Dieser Zustand entspricht einer negativen Spintemperatur und bedeutet eine Polarisation, d. h. bestimmte Ausrichtung der Kernspinverteilung, man spricht von *dynamischer Kernpolarisation.* Die polarisierten Kerne werden u. a. in der Kernphysik verwendet, um bei bestimmten Kernreaktionen die Spin-Spin-Wechselwirkungen gezielt zu untersuchen.

4.4.5.5. Festkörpereffekt. Ähnlich wie der Overhausereffekt wirkt der Festkörpereffekt. In diamagnetischen Kristallen mit paramagnetischer Dotierung strahlt man Mikrowellen mit einer Frequenz $\omega_e + \omega_n$ ein und sättigt die zugehörigen Übergänge. Man erhält ein erheblich größeres Kernresonanzsignal bzw. eine Erhöhung der Kernpolarisation um den Faktor 10^3.

4.4.5.6. Optisches Pumpen. Eine optisch erregte Doppelresonanz liegt beim optischen Pumpen vor. Das Verfahren ist analog zum 3-Niveau-Laser (↑ 4.1.4.). Wie Abb. 4.4.5-2 zeigt, erhöht man durch Einstrahlen von Photonen der Energie $h\nu$ die Besetzungszahl des Niveaus B, das sich z. B. durch Abstrahlung energieärmerer Photonen $h\nu'$ auf das Niveau C abregt. Dieses ist metastabil und entleert sich normalerweise langsam über schwache Relaxationsprozesse.

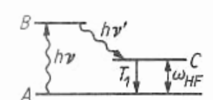


Abb. 4.4.5-2 Optisches Pumpen; Niveauschema und Übergänge

Ist $C \rightarrow A$ ein Hyperfeinstrukturübergang, so erhält man induzierte Mikrowellenemission. Für einen Zeemanübergang (↑ 4.2.4.6.) ergeben sich Hochfrequenzspektren analog zur Kernresonanz. Verwendet man zur Anregung zirkularpolarisiertes Licht, so erzeugt man eine Polarisation des Spinsystems mit einer Vorzugsmagnetisierung, d. h. beispielsweise dynamische Kernpolarisation. Strahlt man zusätzlich ein Hochfrequenzfeld ein, baut sich die Polarisation ab. Damit steigt die Absorption von Photonen. Diese Absorption wird zum Nachweis der optischen Doppelresonanz verwendet.

4.4.5.7. Akustische Doppelresonanz. Als akustische Doppelresonanz ist im Prinzip jede akustische Elektronen- oder Kernresonanz zu betrachten, wenn der Nachweis der akustischen Resonanzabsorption nicht durch reine Schalldämpfungsmessungen, sondern durch paramagnetische Elektronen- bzw. Kernresonanz erfolgt.

Weitere Möglichkeiten sind Kern-Kern- und Elektron-Kern-Doppelresonanzen, bei denen an Stelle des einen elektromagnetischen Hochfrequenzfelds Ultraschall eingestrahlt wird. Der Vorteil gegenüber der rein paramagnetischen Resonanz liegt wieder darin, daß verbotene Übergänge, z. B. mit $|\Delta M| = 2$, erzeugt werden können.

4.4.5.8. Chemisch induzierte Polarisation. Mit den Doppelresonanzen eng verwandt sind die chemisch induzierte Elektronenspinpolarisation und die chemisch induzierte dynamische Kernpolarisation. In reagierenden chemischen Substanzen fand man Änderungen der Spinresonanzsignale, ein Teil der Linien zeigte erhöhte Absorption, andere wandelten sich in Emissionslinien um. Die Ursache hierfür liegt in der Übertragung zusätzlicher Energien in das Spinsystem bei der Entstehung neuer Radikale, z. B. bei Radiolyse. Zur theoretischen Deutung dieser Experimente vermutet man spinselektive chemische Reaktionen.

4.4.6. Kernquadrupolresonanz

Eng verwandt mit der paramagnetischen Kernresonanz ist die Kernquadrupolresonanz. Atomkerne mit einem Kernspin $I > 1/2$ weisen neben dem magnetischen Kerndipolmoment oft auch ein elektrisches Kernquadrupolmoment auf, so daß sich die nicht kugelförmigen Kerne in einem inhomogenen elektrischen Feld ausrichten und um die Richtung des Feldgradienten genauso präzedieren wie im Falle der Kernresonanz das Kerndipolmoment um die Feldrichtung. Das inhomogene elektrische Feld resultiert aus den überlagerten Feldern der Atome bzw. Moleküle der Probensubstanz. Strahlt man ein hochfrequentes elektromagnetisches Feld mit der Präzessionsfrequenz ein, werden Übergänge zwischen den Energieniveaus induziert, die bei der Wechselwirkung zwischen dem Quadrupolmoment und dem inhomogenen Feld entstehen, so daß die Kerne umorientiert werden. Die dabei auftretende Energieabsorption kann man wie bei der Spinresonanz nachweisen und auswerten. Die benutzten Frequenzen entsprechen Wellenlängen im Meter- und Dezimeterbereich. Charakteristisch für die Quadrupolresonanz ist das Fehlen eines konstanten Felds, der Nachweis der Resonanz ist daher nur über Frequenzmodulation möglich, d. h., man ändert die Frequenz des Hochfrequenzfelds kontinuierlich und erhält so die Absorptionslinien. Mit der Kernquadrupolresonanz lassen sich Informationen über die Kernquadrupolmomente und die Feldgradienten im Inneren der Probe gewinnen und damit Aussagen über die Struktur des Kristallgitters, über Gitterstörungen sowie über chemische Bindungen machen. Hauptanwendungsgebiet ist die Untersuchung kovalenter Bindungen, besonders an Halogeniden.

5. Moleküle

5.1. Chemische Bindung

Moleküle entstehen durch Vereinigung zweier oder mehrerer gleicher bzw. verschiedener Atome. Ihre Energie liegt unter der Summe der Energien der einzelnen Atome; der Differenzbetrag wird bei der Molekülbildung freigesetzt. Moleküle enthalten jeweils nur ganz bestimmte Atome in bestimmten Atomkombinationen. Zum Beispiel vereinigen sich Wasserstoffatome im Gaszustand zu H_2- und nicht zu H_3- oder H_4-Molekülen, andererseits kombinieren zwei Heliumatome nicht unter Bildung eines He_2-Moleküls. Eine Erklärung für den Energiegewinn bei der Molekülbildung sowie für das Auftreten bzw. Ausbleiben der jeweiligen Atomkombinationen vermag, zumindest in einfachen Fällen, die Theorie der chemischen Bindung zu geben.

5.1.1. Klassische Vorstellungen zur Beschreibung der chemischen Bindung

5.1.1.1. Entwicklung der Valenztheorie. Obwohl die Vorstellung vom Bestehen einer chemischen Bindung zwischen zwei Atomen A und B, in der chemischen Symbolsprache z. B. mit A–B bezeichnet, sehr alt ist, erfolgte ein Durchbruch bei der theoretischen Erklärung dieses Sachverhalts erst im 20. Jahrhundert. Über die wirkliche Beschaffenheit der Bindung herrschte zuvor völlige Unklarheit. So hat noch GASSENDI in seinem 1658 entstandenem Werk „Syntagma Philosophicum" ähnlich wie die frühen griechischen Atomisten vorgeschlagen, daß sich Atome mit Hilfe von Häkchen an ihrer Oberfläche aneinanderheften könnten. Noch Mitte des 19. Jahrhunderts war das Bild eines Atoms ähnlich dem einer Billardkugel, die zusätzlich auf ihrer Oberfläche eine bestimmte Anzahl von „Knöpfen" zur Verbindung der Atome miteinander aufwies. 1852 schlug FRANKLAND vor, durch den Begriff *Wertigkeit* oder *Valenz* auszudrücken, daß sich die Elemente immer nur in ganz bestimmten Mengeneinheiten miteinander verbinden. Die Wertigkeit eines Elements ist danach gleich der Zahl der Wasserstoff- oder ihm äquivalenter Atome, die gebunden oder ersetzt werden können.

1857/58 schrieben KEKULÉ, KOLBE und COUPER dem Kohlenstoff die Wertigkeit vier zu. Noch weitergehende Aussagen, wie die 1874 von VAN'T HOFF und LE BEL vorgeschlagene tetraedrische Anordnung der vier Wasserstoffatome des Methans um das zentrale Kohlenstoffatom, wurden von den bedeutenden Experimentalchemikern dieser Zeit als spekulativ, ja als prinzipiell unbeweisbar abgelehnt.

Auf die Entdeckung des Elektrons (↑ 1.1.2.1.) folgten zahlreiche Versuche, eine Elektronentheorie der chemischen Bindung zu entwickeln. Die Bemühungen gipfelten in den 1916 veröffentlichten Arbeiten von LEWIS und KOSSEL, wonach die Atome bestrebt seien, bei der Verbindungsbildung eine *Edelgaskonfiguration* ihrer Elektronen zu erhalten.

Die Verfeinerung der Elektronentheorie der Bindung bis zur gegenwärtigen Gestalt gelang vorwiegend mit Hilfe der *Quantenmechanik*. Außer der Zahl und Anordnung der Bindungen lassen sich inzwischen für viele einfachere Moleküle – zumindest näherungsweise – die Bindungsenergien und andere Bindungsparameter quantitativ berechnen.

5.1.1.2. Grenztypen der chemischen Bindung. Die Atome der reinen Elemente werden meist durch kovalente oder metallische Bindungen zusammengehalten, zumindest aber durch unspezifische Wechselwirkung vom van-der-Waals-Typ (↑ 5.1.4.1.). Im allgemeinen ist der Bindungscharakter in solchen homonuklearen Molekülen oder Baugruppen nicht durch einen dieser Grenztypen allein gegeben, sondern liegt zwischen ihnen.

Beim Zusammentreten unterschiedlicher Elemente zu nunmehr heteronuklearen Kombinationen, den chemischen Verbindungen, ergibt sich aufgrund der unterschiedlichen Elektronegativität der Elemente die Ionenbeziehung als weiterer Grenztyp einer chemischen Bindung.

Bei der klassischen Beschreibung des Phänomens der chemischen Bindung spielt die Stabilität der Elektronenkonfiguration der Edelgase eine zentrale Rolle. Die genannten Bindungstypen lassen sich dementsprechend vor allem durch die unterschiedlichen Wege charakterisieren, auf denen die Bindungspartner ihre Edelgaskonfiguration erreichen oder ihr doch nahekommen. Auch nach Erarbeitung des wellenmechanischen Atommodells und der daraus abgeleiteten Bindungstheorie, die um 1930 im wesentlichen vorlag, spielten die Vorstellungen von LEWIS und KOSSEL noch lange Zeit im chemischen Schrifttum eine dominie-

rende Rolle bei der Diskussion von Bindungsproblemen. Erst nach Entdeckung der ersten Edelgasverbindungen (↑ 5.3.2.1.) im Jahre 1962 und der Erkenntnis, daß die Bindungsverhältnisse in diesen Verbindungen durchaus normal sind, kam es zum umfassenden Bruch mit der Vorstellung von der inerten Edelgaskonfiguration.

5.1.1.2.1. *Metallische Bindung.* Sie ist der charakteristische Bindungstyp in den Kristallgittern der Metalle, d. h. bei der Mehrzahl der Elemente.

Nach einer von LORENTZ aufgestellten Theorie bestehen die Metalle aus regelmäßig angeordneten Atomrümpfen, zwischen denen sich die abgespalteten Valenzelektronen als *Elektronengas* frei bewegen. Die Anziehung zwischen positiv geladenen Atomrümpfen und negativ geladenen Elektronen bewirkt den Zusammenhalt des Gitters.

Bei den Nebengruppenelementen werden weniger Elektronen abgespalten, als zum Erreichen der Elektronenkonfiguration des jeweils zuvorstehenden Edelgases notwendig wäre. Immerhin weisen die Valenzelektronenkonfigurationen in mehreren Fällen, z. B. bei Chromium, Kupfer, Gadolinium und Curium, eine relative Bevorzugung halb- oder vollbesetzter *d*- oder *f*-Elektronenzustände aus, die mit den eigentlichen Edelgaskonfigurationen die sphärische Symmetrie der Ladungsverteilung gemeinsam haben.

Das Elektronengasmodell kann zwar einige typische Metalleigenschaften, wie die elektrische bzw. thermische Leitfähigkeit sowie den metallischen Glanz, qualitativ gut erklären. Eine umfassendere Beschreibung des metallischen Zustands kann aber nur mit Hilfe quantenmechanischer Vorstellungen erfolgen.

5.1.1.2.2. *Atombindung oder Kovalenz.* Eine weitere Möglichkeit zur Erlangung einer Edelgaskonfiguration besteht in der gemeinsamen Beanspruchung eines oder mehrerer Elektronenpaare durch zwei Atome. So erreichen im Wasserstoffmolekül, H: H bzw. H—H, beide H-Atome dadurch die Elektronenkonfiguration des Edelgases Helium, daß das bindende Elektronenpaar beiden Atomen gemeinsam angehört. Der Valenzstrich zwischen den Atomen stellt jeweils ein gemeinsames Elektronenpaar dar; er symbolisiert somit mehr als den bloßen Zusammenhalt der Teilchen.

Der Zusammenhalt des Moleküls beruht auf der elektrostatischen Wechselwirkung zwischen dem gemeinsamen Elektronenpaar und den beiden Atomkernen.

Im Sauerstoffmolekül, $\ddot{\underset{..}{O}}::\ddot{\underset{..}{O}}$ bzw. $\overline{\underline{O}}=\overline{\underline{O}}$, ist die Edelgaskonfiguration durch zwei und im Stickstoffmolekül $|N{\equiv}N|$ durch drei gemeinsame Elektronenpaare verwirklicht. Manche Elemente haben zu wenig Elektronen in der Valenzschale, um bei der Bildung neutraler Moleküle Edelgaskonfiguration zu erreichen. Das Bor hat z. B. im BF_3-Molekül nur sechs Elektronen in seiner äußeren mit Elektronen besetzten Schale. Durch Anlagerung eines Fluoridions entsteht das Tetrafluoroboration mit Oktettkonfiguration des Bors:

$$BF_3 + F^{\ominus} \rightarrow \left[\begin{array}{ccc} & F & \\ & | & \\ F— & B & —F \\ & | & \\ & F & \end{array}\right]^{\ominus}.$$

Das Bor bestätigt im $[BF_4]^-$-Ion vier kovalente Bindungen. Unter der Annahme, daß sich die beiden Elektronen eines bindenden Elektronenpaars gleichmäßig auf die zwei verbundenen Atome verteilen, kommt dem Bor im $[BF_4]^-$-Ion ein Elektron mehr als im neutralen Atom zu. Daraus resultiert die Ladung -1. Sie wird als *Formalladung* bezeichnet und häufig durch ein entsprechendes Zeichen über dem Atomsymbol dargestellt: $[\overset{\ominus}{B}F_4]^-$. Sie darf nicht mit der Oxydationszahl oder -stufe eines Atoms in einem Molekül verwechselt werden. Zur Ermittlung der *Oxydationszahl* denkt man sich ein Molekül so in Ionen zerlegt, daß die Bindungselektronenpaare dem jeweils elektronegativeren Bindungspartner vollständig zugeordnet werden (↑ 5.1.3.3). Bei Bindungen zwischen gleichartigen Atomen werden die Elektronenpaare halbiert. Die Oxydationszahl ergibt sich dann aus der Differenz zwischen der Elektronenzahl der resultierenden Ionen und der der entsprechenden neutralen Atome. Letztere haben definitionsgemäß die Oxydationszahl Null. Um Verwechslungen mit der Ionenwertigkeit zu vermeiden, schreibt man im Bedarfsfalle Oxydationszahlen mit vorgestelltem Vorzeichen über das Elementsymbol. Die Summe der Oxydationszahlen aller Atome eines Moleküls oder Ions muß der Ladung der Gesamtpartikel entsprechen. Für Bor im $[BF_4]^-$-Ion ergibt sich z. B. die Oxydationszahl $+3$, für Fluor -1, $[\overset{+3}{B}(\overset{-1}{F})_4]^-$. Schwefel hat im Sulfation die Oxydationszahl $+6$, Sauerstoff -2. Obwohl Oxydationszahlen ebenfalls nur formalen Charakter haben, sind sie ein wichtiges Hilfsmittel bei der Formulierung von Redoxgleichungen sowie für die Abschätzung möglicher Zwischen- oder Endprodukte chemischer Reaktionen (↑ 7.2.).

5.1.1.2.3. *Ionenbeziehung oder Elektrovalenz.* Eine dritte Möglichkeit zur Ausbildung einer Edelgaskonfiguration besteht in der Aufnahme bzw. Abgabe von Elektronen unter Ausbildung von Atomionen. Treten z. B. ein Natrium- und ein Chloratom in der Gasphase in Wechselwirkung, so nimmt das Natrium unter Abgabe seines *Valenzelektrons* die Oktettkonfiguration

des Neons an, während das Chloratom unter Aufnahme dieses Elektrons die Konfiguration des Edelgases Argon erreicht. Zur Bildung des Ions Na^+ muß die Ionisierungsenergie des Natriums aufgewendet werden, andererseits wird bei der Bildung des Ions Cl^- ein als *Elektronenaffinität* des Chloratoms bezeichneter Energiebetrag freigesetzt.

Die *Ionisierungsenergie* oder das 1. *Ionisierungspotential* ist die Mindestenergie, die zur Entfernung eines Elektrons aus einem Atom, Ion oder Molekül aufgewendet werden muß. Die *Elektronenaffinität* ist die beim Einbau eines zusätzlichen Elektrons in ein Atom, Ion oder Molekül freiwerdende Energie.

Ionisierungsenergien haben ein positives Vorzeichen. Die Vorzeichenregelung für Elektronenaffinitäten ist der für Ionisierungsenergien entgegengesetzt, d. h., sie erhalten bei freiwerdender Energie ein positives Vorzeichen. Negative Elektronenaffinitäten zeigen also an, daß zur Elektronenaufnahme Energie aufgewendet werden muß; das ist z. B. bei der Ausbildung zweifach negativ geladener Ionen aus einfach negativ geladenen der Fall.

Die Ionisierungsenergien erreichen bei den Edelgasen bzw. den Atomionen mit gleicher Elektronenzahl ein Maximum, die Elektronenaffinitäten hingegen ein Minimum (Abb. 5.1.1-1). Hierin drückt sich die besondere Stabilität der Edelgaskonfiguration sowie die Tatsache aus, daß die höchstmögliche *Ionenwertigkeit* der Atomionen in Verbindungen, d. h. deren maximale Ladungszahl, durch die Entfernung des betreffenden Atoms zum nächsten Edelgas bestimmt wird.

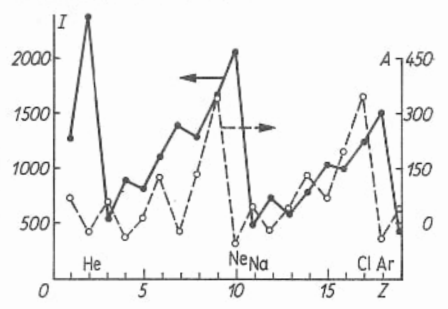


Abb. 5.1.1-1 Ionisierungspotentiale I und Elektronenaffinitäten A in $kJmol^{-1}$ für einige Hauptgruppenelemente in Abhängigkeit von deren Ordnungszahl Z

Verbindungen mit vorwiegend ionischem Bindungscharakter sind immer dann zu erwarten, wenn Wechselwirkungen zwischen Atomen mit relativ niedrigem Ionisierungspotential und solchen mit großer Elektronenaffinität erfolgen.

Der Vergleich der Ionisierungsenergie von Na mit der Elektronenaffinität von Cl zeigt (↑ Abb. 5.1.1-1), daß die Bildung von Na^+- und Cl^--Ionen aus den Atomen unter Energieaufwand erfolgt. Selbst für die Bildung von Cs^+F^- resultiert eine ungünstige Energiebilanz. Dies spräche eher für einen kovalenten als für einen ionischen Aufbau der NaCl-Moleküle. Bei Annäherung entgegengesetzt geladener Ionen auf den Gleichgewichtsabstand im Molekül sinkt jedoch die potentielle Energie nach dem Coulombschen Gesetz beträchtlich ab. Der ionische Zustand kann folglich beim Gleichgewichtsabstand der Bindungspartner einen niedrigeren Energieinhalt als der Atomzustand haben. Ein noch höherer Energiebetrag wird bei der Bildung eines NaCl-Kristalls freigesetzt, in dem jedes Na^+-Ion von sechs Cl^--Ionen umgeben ist und umgekehrt. Die Ionenbeziehung hat demzufolge vorwiegend Bedeutung als Bindungstyp bei kristallinen Verbindungen (↑ 6.4.3.1.).

Der Begriff Ionenbeziehung anstelle von Ionenbindung trägt der Tatsache Rechnung, daß die Coulombwechselwirkungen zwischen den Ionen im Festkörper nicht auf Vorzugsrichtungen lokalisiert sind, von Bindungen im eigentlichen Sinne also nicht gesprochen werden kann.

Während in homonuklearen Molekülen rein kovalente Bindungen auftreten können, sind rein ionische Bindungen aufgrund der Verhältnisse zwischen Ionisierungspotentialen und Elektronenaffinitäten in keinem Falle realisiert (↑ 5.1.3).

5.1.2. Quantenmechanische Beschreibung der chemischen Bindung

Das Hauptproblem einer Theorie der chemischen Bindung besteht darin, die Verteilung der Elektronendichte im Molekül zu bestimmen. Hierzu ist es erforderlich, die Schrödingergleichung (↑ 2.5.3.) $\hat{H}\psi = E\psi$ für das betrachtete Molekül zu lösen (lies: $\hat{H}$ auf psi gleich E mal psi). In den Hamiltonoperator $\hat{H}$ gehen die Koordinaten der Kerne und der Elektronen ein. Die Kerne bewegen sich wegen ihrer relativ hohen Masse viel langsamer im Molekül als die Elektronen. Nach Born und Oppenheimer ist es deshalb gerechtfertigt, Kern- und Elektronenbewegung im Molekül unabhängig voneinander zu behandeln. Im Rahmen dieser *adiabatischen Näherung* reduziert sich also unser Problem auf eine Lösung der Schrödingergleichung für stationäre Zustände der Elektronen im Feld des starren Kerngerüsts. Die Schrödingergleichung ist aber auch dann nur für das einfachste Molekül, das H_2^+-Ion, exakt lösbar. Zur quantenmechanischen Behandlung größerer Moleküle ist man zu weiteren Näherungsannahmen gezwungen.

5.1.2.1. Wasserstoffmolekülion H_2^+. Das einfachste Molekül ist aus zwei Protonen und einem Elektron aufgebaut (Abb. 5.1.2-1); es wurde von THOMPSON in Katodenstrahlen entdeckt. Seine Existenz kann durch nichtquantenmechanische Theorien, die vom Konzept einer Zweielektronenbindung ausgehen, nicht erklärt werden. Die quantenmechanische Behandlung des H_2^+-Problems hat zentrale Bedeutung, da sie die Grundlage einer Näherungstheorie für komplexere zweiatomige Moleküle ist.

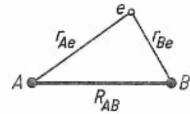


Abb. 5.1.2-1 Größen zur Beschreibung des H_2^+-Molekülions

Im Rahmen der Born-Oppenheimer-Näherung ist das H_2^+-Problem exakt lösbar. Der Hamiltonoperator $\hat{H}$ (↑ 2.5.2.2.) hat die Form

$$(1) \quad \hat{H} = -h^2/(8\pi^2 m\Delta) + e^2/(4\pi\varepsilon_0 R_{AB}) - e^2/(4\pi\varepsilon_0 r_{Ae}) - e^2/(4\pi\varepsilon_0 r_{Be}).$$

An dieser Stelle soll das *Variationsverfahren* zur Lösung skizziert werden. Es wird dabei von einer Näherungsfunktion mit mehreren willkürlichen Parametern ausgegangen, die die Randbedingungen des Problems erfüllt. Bei der Variationsberechnung des H_2^+ kann eine Linearkombination der atomaren 1s-Orbitale Φ_A und Φ_B des Wasserstoffs zugrunde gelegt werden: $\psi = c_A\Phi_A + c_B\Phi_B$. Es werden nun die Parameter c_A und c_B bestimmt, bei denen sich ein Minimum der Energie ergibt.

Durch Multiplikation der Schrödingergleichung von links mit ψ und Integration über den gesamten Raum erhält man für die Energie die Beziehung:

$$(2) \quad E = \frac{\int \psi \hat{H} \psi \, d\tau}{\int \psi^2 \, d\tau} = \frac{c_A^2\alpha_A + 2c_Ac_B\beta_{AB} + c_B^2\alpha_B}{c_A^2 S_{AA} + 2c_Ac_B S_{AB} + c_B^2 S_{BB}}$$

mit den Abkürzungen $\int \Phi_A \hat{H} \Phi_B \, d\tau = \beta_{AB}$ = *Resonanzintegral*, $\int \Phi_A \hat{H} \Phi_A \, d\tau = \alpha_A$ = *Coulombintegral* und $\int \Phi_A \Phi_B \, d\tau = S_{AB}$ = *Überlappungsintegral*. Zur Ermittlung der Parameter c_A und c_B, für die E ein Minimum ist, müssen die partiellen Ableitungen $\partial E/\partial c_A$ und $\partial E/\partial c_B$ gebildet und gleich Null gesetzt werden. Nach einigen algebraischen Umformungen führt das auf die Gleichungen $c_A(\alpha_A - ES_{AA}) + c_B(\beta_{AB} + ES_{AB}) = 0$ und $c_A(\beta_{AB} - ES_{AB}) + c_B(\alpha B - ES_{BB}) = 0$. Dieses System homogener linearer Gleichungen, die *Säkulargleichungen* genannt werden, hat nur dann nichttriviale, d. h. andere Lösungen als $c_A = c_B = 0$, wenn die Determinate der Koeffizienten, die *Säkulardeterminante*, verschwindet. Unter Berücksichtigung der Normiertheit der atomaren Funktionen, d. h. $S_{AA} = S_{BB} = 1$ und der Gültigkeit von $\alpha_A = \alpha_B = \alpha$ beim H_2^+-Problem ergibt sich

$$\begin{vmatrix} (\alpha - E) & (\beta_{AB} - ES_{AB}) \\ (\beta_{AB} - ES_{AB}) & (\alpha - E) \end{vmatrix} = 0.$$

Es resultieren die beiden Lösungen

$$(3) \quad E_+ = \frac{\alpha + \beta_{AB}}{1 + S_{AB}} \quad \text{und}$$

$$(4) \quad E_- = \frac{\alpha - \beta_{AB}}{1 - S_{AB}}.$$

Einsetzen von E_+ und E_- in die Säkulargleichungen und Berücksichtigen der Normierungsbedingung $\int \psi^2 \, d\tau = 1$ liefern die Koeffizienten $c_{A^+} = c_{B^+} = \frac{1}{\sqrt{2(1 + S_{AB})}}$ und $c_{A^-} = -c_{B^-} = \frac{1}{\sqrt{2(1 - S_{AB})}}$. Die den Energien E_+ und E_- zugeordneten Wellenfunktionen lauten

$$(5) \quad \psi_+ = c_A\Phi_A + c_A\Phi_B = (2 + 2S_{AB})^{-1/2}(\Phi_A + \Phi_B)$$

$$(6) \quad \psi_- = c_A\Phi_A - c_A\Phi_B = (2 - 2S_{AB})^{-1/2}(\Phi_A - \Phi_B).$$

Die Integrale α, β_{AB} und S_{AB} lassen sich nach Transformation in elliptische Koordinaten lösen.

Bei Annäherung der Atome in den Bereich des Kernabstands R_0 im H_2^+ hat das Coulombintegral α negative Werte, die annähernd der Energie eines 1s-Orbitals des Wasserstoffs entsprechen; das Resonanzintegral β_{AB} und das Überlappungsintegral S_{AB}, die bei großem Kernabstand den Wert Null haben, werden zunehmend negativ (β_{AB}) bzw. positiv (S_{AB}). Dadurch

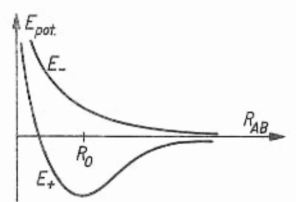


Abb. 5.1.2-2 Energien der Zustände ψ_+ und ψ_- von H_2^+ als Funktion des Kernabstands R_{AB}

wird die in Abb. 5.1.2-2 schematisch dargestellte Variation der Energien E_+ und E_- als Funktion des Kernabstands verständlich.

Als Ergebnis der näherungsweisen Berechnung des H_2^+ kann also festgestellt werden, daß aus der Wechselwirkung der beiden 1s-Atomorbitale von Wasserstoff zwei Einelektronenwellen-

funktionen des Moleküls ψ_+ und ψ_- resultieren, von denen im Bereich des Gleichgewichtsabstands ψ_+ eine kleinere ψ_- eine größere Energie als der 1*s*-Zustand des Wasserstoffatoms haben. ψ_+ wird als bindendes und ψ_- als antibindendes Molekülorbital oder MO bezeichnet. Befindet sich das Elektron im Zustand ψ_+, so liegt ein gegenüber H + H^+ energieärmerer Elektronenzustand vor. Dieser Zustand hat im Gleichgewichtsabstand R_0 seine geringste Energie. Die Theorie führt somit zur Vorhersage, daß aus einem H-Atom und einem Proton ein stabiles Molekülion H_2^+ gebildet wird.

Die Bildung von Molekülorbitalen durch Linearkombination von Atomorbitalen, d. h. die *LCAO-MO-Methode*, ist auch auf komplexe Moleküle übertragbar. Bei einer LCAO-Funktion mit n Atomorbitalen resultiert ein Satz von n Säkulargleichungen mit jeweils n Koeffizienten. Die Lösung der $n \times n$-Säkulardeterminante führt zu einem Polynom n-ten Grades in E. Die n Wurzeln des Polynoms entsprechen den Energien der einzelnen Molekülorbitale. Einsetzen der errechneten Energien in die Säkulargleichungen liefert die Koeffizienten der jeweiligen LCAO-Funktionen.

Molekülorbitale sind Einelektronenwellenfunktionen. Die Gesamtwellenfunktion des Moleküls wird bei den Atomen als Produkt der Einelektronenwellenfunktionen unter Berücksichtigung der Antisymmetriebedingung aufgebaut.

5.1.2.2. Überlappung und Bindungsstärke. Nach Gleichung 5.1.(3) ist die Bindung in einem Molekül um so stärker, je größer der Wert des Überlappungsintegrals und der Betrag des negativen Resonanzintegrals ist. Beide Größen sind proportional, wobei das Resonanzintegral den größten Beitrag für die Energieabsenkung des bindenden Molekülorbitals liefert. Aus dieser Überlegung ist ein einfaches Kriterium für die Stärke einer Bindung ableitbar: Je stärker die Überlappung zwischen den Atomorbitalen ist, um so stärker ist die chemische Bindung.

In analoger Weise wie 1*s*-Orbitale lassen sich auch andere Atomorbitale zu Molekülorbitalen kombinieren. Dabei ergibt sich die Art der Überlappung aus der der Symmetrie der Bahnfunktionen. So läßt sich z. B. für die in Abb. 5.1.2-3 dargestellte Kombination eines *s*- mit einem *p*-Orbital für das Überlappungsintegral und damit auch für das Resonanzintegral ein Wert von genau Null voraussagen, da sich die Anteile positiver und negativer Überlappung aufheben. Eine derartige Kombination führt zu sogenannten nichtbindenden Zuständen.

Die bei der Kombination zweier *s*-Orbitale resultierenden Molekülorbitale sind rotationssymmetrisch zur Bindungsachse. Solche Molekülorbitale werden als σ-Orbitale, die entsprechenden Bindungen als σ-Bindungen bezeichnet; die jeweils antibindenden Zustände werden zusätzlich durch einen Stern markiert. Positive Überlappung führt zu einem bindenden σ-MO, negative Überlappung zu einem antibindenden σ^*-MO. Bezeichnet man in einem zweiatomigen Molekül die Kernverbindungsachse als z-Achse eines kartesischen Koordinatensystems, ergibt sich, daß die beiden p_z-Orbitale der beteiligten Atome ebenfalls unter Ausbildung einer σ-Bindung überlappen. Die verbleibenden p_x- und p_y-Orbitale können nur zu bindenden π- und

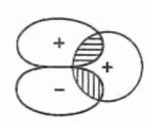


Abb. 5.1.2-3 Überlappung von Atomorbitalen zu einem nichtbindenden Molekülorbital

antibindenden π^*-MO überlappen; π-Orbitale haben eine Knotenebene in der Kernverbindungsachse (Abb. 5.1.2-4). Die für Bindungsprobleme wenig bedeutsamen δ-Orbitale mit zwei senkrecht aufeinander stehenden Knotenebenen in der Kernverbindungsachse resultieren aus der Überlappung von $2d_{xy}$-Atomorbitalen.

5.1.2.3. Molekülorbitalbehandlung homonuklearer zweiatomiger Moleküle. Zur Beschreibung der Bindung in homonuklearen zweiatomigen Molekülen müßte die Wechselwirkung zwischen allen Orbitalen mit gleichem Symmetrieverhalten berücksichtigt werden. Für qualitative Betrachtungen ist es jedoch ausreichend, nur die Überlappung zwischen Atomorbitalen gleicher Energie zu berücksichtigen, da die Wechsel-

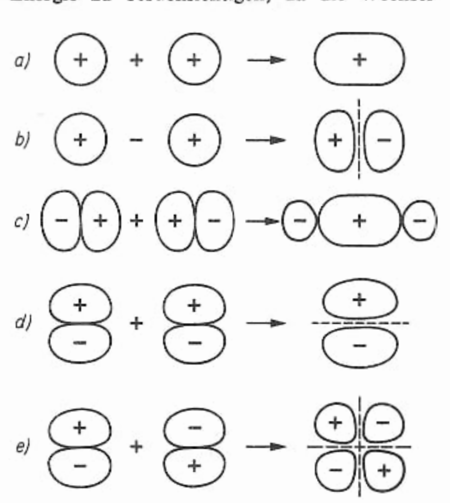


Abb. 5.1.2-4 Kombination von Atomorbitalen zu Molekülorbitalen verschiedener Typen; a) σ_s, b) σ_s^*, c) σ_p, d) π_p, e) π_p^*

wirkung mit wachsender Energiedifferenz der zu kombinierenden Zustände stark abnimmt. Unter dieser Annahme ist für zweiatomige Moleküle der zweiten Periode die Wechselwirkung zwischen jeweils gleichen s- oder p-Orbitalen zu untersuchen. Abb. 5.1.2-5 veranschaulicht das

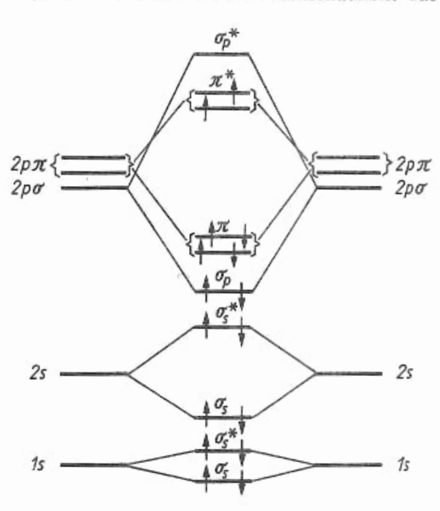


Abb. 5.1.2-5 Qualitatives Molekülorbitalschema für ein zweiatomiges Molekül am Beispiel von O_2

resultierende Molekülorbitalschema. Bei Atomen der 1. Achterperiode wird mit wachsender Ordnungszahl die Wechselwirkung zwischen den $1s$-Funktionen wegen der kaum noch zustandekommenden Überlappung immer geringer; die Energiedifferenz zwischen σ_{1s} und σ_{1s}^* ist demzufolge klein gegenüber der zwischen σ_{2s} und σ_{2s}^*. Die in den freien Atomen gleichwertigen, also *energetisch entarteten* $2p$-Orbitale kombinieren in unterschiedlicher Weise, entsprechend ihrer jeweiligen Orientierung zur Kernverbindungsachse, zu Molekülorbitalen. Legt man wieder die Kernverbindungsachse als z-Richtung fest, so kombinieren die p_z-Orbitale zu σ-MO, die p_x- und p_y-Orbitale dagegen zu π- und π^*-MO, die bis auf eine 90°-Verdrehung um die Kernverbindungsachse äquivalent und weiterhin entartet sind. Im Rahmen der Theorie unabhängiger Elektronen, d. h. unter Vernachlässigung der Elektronenwechselwirkung, werden die bindenden und antibindenden Molekülorbitale analog dem Aufbauprinzip der Atome in der Reihenfolge zunehmender Energien mit Elektronen besetzt. Dabei muß das Pauliprinzip berücksichtigt werden. Es resultiert die *Elektronenkonfiguration* des Moleküls im Grundzustand. Die jeweilige *Bindungsordnung* ergibt sich als Differenz der Zahl der Elektronenpaare in bindenden und antibindenden MO (↑ 5.1.2.1.). Beim H_2-Molekül ist die Bahnfunktion niedrigster Energie mit zwei Elektronen besetzt, die entsprechend dem Pauliprinzip antiparallelen Spin haben. Bei der Kombination zweier He-Atome sind in das MO-Schema vier Elektronen einzubauen. Es resultiert die Elektronenkonfiguration $(\sigma_{1s})^2(\sigma_{1s}^*)^2$. Während beim H_2-Molekül die Bindungsordnung $(1-0) = 1$ beträgt, ist sie beim He_2 $(1-1) = 0$, d. h. im Einklang mit dem atomaren Auftreten des Edelgases Helium. Zur Ermittlung der Elektronenkonfiguration des O_2-Moleküls sind in das MO-Schema 16 Elektronen einzubauen. Es resultiert die Verteilung

$(\sigma_{1s})^2\ (\sigma_{1s}^*)^2\ (\sigma_{2s})^2\ (\sigma_{2s}^*)^2\ (\sigma_{2p})^2\ (\pi_{2p})^4\ (\pi_{2p}^*)^2$.

Die Bindungsordnung beträgt $(1 + 1 + 1 + 2) - (1 + 1 + 1) = 2$. Die Besetzung des π^*-Niveaus erfolgt gemäß der *Hundschen Regel* (↑ 4.2.4.3.) mit parallelem Spin der beiden Elektronen. Dies ist die Ursache für den experimentell beobachteten *Paramagnetismus* des Sauerstoffs, der auf der Grundlage klassischer Bindungsvorstellungen nicht verständlich war.

5.1.2.4. Lokalisierte Molekülorbitale und Hybridisierung. Auch die Molekülorbitale komplizierter Moleküle können durch Linearkombination der Atomorbitale der beteiligten Atome erhalten werden. Man kann zum Verständnis der Moleküleigenschaften gelangen, ohne vom klassischen Konzept lokalisierter Bindungen auszugehen. Dieses Konzept ist dennoch nützlich für die Abschätzung von Moleküleigenschaften; es findet seine Rechtfertigung in zahlreichen experimentellen Tatsachen. IR-Frequenzen bestimmter Bindungen sind z. B. weitgehend unabhängig vom jeweiligen Molekülrest, was nicht zuletzt die Grundlage für die Anwendung der IR-Spektroskopie zur Identifizierung unbekannter Verbindungen ist (↑ 5.3.4.). Es läßt sich auch quantenmechanisch durch die Annahme rechtfertigen, daß ein Teil der zu betrachtenden Wechselwirkungen stark, der andere Teil schwach ist.

Die Struktur zahlreicher drei- und mehratomiger Moleküle läßt sich verstehen, wenn man lokalisierte Molekülorbital ezugrunde legt, die aus der Wechselwirkung von Atomorbitalen je zweier Atome resultieren, wobei die Atomorbitale jeweils selbst schon Linearkombinationen *verschiedener* Atomfunktionen *eines* Atoms sein können. Das soll am Beispiel des Methanmoleküls CH_4 näher erläutert werden: Unter der Annahme, daß C–H-Bindungen durch Überlappung halbbesetzter Atomorbitale von Wasserstoff und Kohlenstoff entstehen, würde man für Kohlenstoff infolge der Elektronenkonfiguration seines Grundzustands (Abb. 5.1.2-6) die Wertigkeit zwei und damit die Existenz eines stabilen Moleküls „CH_2" erwarten. Die beobachtete Vierwertigkeit von

Kohlenstoff im Methan verlangt die Annahme der Existenz eines angeregten Zustands mit vier ungepaarten Valenzelektronen. Legt man dafür den energieärmsten Fall gemäß Abb. 5.1.2-6b zugrunde, sollte man für CH_4 zwei unterschiedliche C—H-Bindungen erwarten. Tatsächlich liegen vier völlig gleichwertige vor; sie erstrekken sich in Richtung der Ecken eines regulären Tetraeders. Eine befriedigende Beschreibung der Bindung gelingt somit erst unter der Annahme, daß sie durch vier völlig gleichwertige *sp^3-Hybridorbitale* am Kohlenstoff vermittelt wird. Sie stellen Linearkombinationen der $2s$- und der drei $2p$-Bahnfunktionen mit jeweils $1/4\,s$- und $3/4\,p$-Charakter dar (Abb. 5.1.2-6c).

Die Energie, die bei der Bildung von vier statt zwei C—H-Bindungen gewonnen wird, ist größer als die zur Anregung von $1s^2 2s^2 2p^2$ zu $1s^2 2s 2p 2p 2p$ und zur Promotion in den sp^3-Valenzzustand erforderliche, so daß die Bildung von CH_4 gegenüber der eines hypothetischen „CH_2" begünstigt ist. Die Bildung des sp^3-Valenzzustands ist kein beobachtbarer physikalischer Vorgang, dennoch eine für die Diskussion der Bindung nützliche Gedankenkonstruktion.

5.1.2.4.1. *Mehrfachbindungen und Bindungswinkel.* Die Struktur der ungesättigten, d. h. C—C-Mehrfachbindungen enthaltenden, Kohlenwasserstoffe läßt sich bei Annahme einer sp^2- bzw. sp-Hybridisierung des Kohlenstoffs zwanglos erklären. Ein Satz von drei sp^2-Hybridorbitalen liegt in einer Ebene, die Winkel zwischen den Orbitalen betragen 120°. Zwei sp-Hybridorbitale sind entsprechend im Winkel von 180° zueinander orientiert.

Ethylen oder Ethen C_2H_4 ist ein ebenes Molekül, alle Bindungswinkel betragen etwa 120°. Das Molekülgerüst besteht aus σ-Bindungen zwischen sp^2-Hybridorbitalen der beiden C-Atome und $1s$-Orbitalen der H-Atome. Die an der Hybridisierung nicht beteiligten p-Orbitale sind senkrecht zur Molekülebene orientiert.

a) 2s 2p
b) 2s 2p
c) 4×sp^3

Abb. 5.1.2-6 Elektronenkonfiguration von Kohlenstoff; a) im Grundzustand, b) in einem angeregten Zustand, c) im hypothetischen sp^3-Valenzzustand

Sie sind jeweils mit einem Elektron besetzt und überlappen unter Ausbildung einer π-Bindung. Die C—C-Doppelbindung besteht demzufolge aus einer σ- und einer π-Bindung (Abb. 5.1.2-7a). Im linearen Acetylen- oder Ethinmolekül C_2H_2 besteht das Bindungsgerüst aus σ-Bindungen zwischen sp-Hybridorbitalen der beiden C-Atome und den $1s$-Orbitalen der H-Atome. Die an der Hybridisierung nicht beteiligten zwei p-Orbitale pro C-Atom sind senkrecht zueinander orientiert und jeweils mit einem Elektron besetzt. Sie kombinieren unter Ausbildung zweier

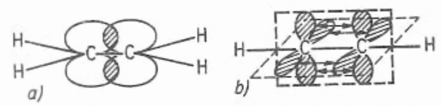


Abb. 5.1.2-7 Bindungsverhältnisse a) bei Ethylen, b) bei Acetylen; das σ-Bindungsgerüst ist durch *Striche* symbolisiert

a) b) c)

Abb. 5.1.2-8 Mesomerie am Beispiel des NO_2^--Ions

ebenfalls senkrecht zueinander orientierter π-Bindungen (Abb. 5.1.2-7b).

5.1.2.5. Valenzstrichformeln und Mesomerie. Der Zusammenhalt der Atome in Molekülen oder Molekülionen enthält durch die Formulierung von *Valenzstrichformeln* mit Angabe der bindenden und freien Valenzelektronenpaare einen anschaulichen Ausdruck. Für HF, Ethan, Formaldehyd, NH_3 bzw. das Ammoniumion NH_4^+ ergeben sich z. B. die Formeln

sie werden auch als *Konstitutionsformeln* bezeichnet und geben die Elektronenverteilung in richtiger Weise wieder. Oftmals ist es jedoch nicht möglich, ein Molekül oder Molekülion durch *eine* Valenzstrichformel befriedigend zu beschreiben. Dies gilt z. B. für das Nitrition (Abb. 5.1.2-8). Das Molekül ist gewinkelt, die beiden N—O-Bindungslängen sind identisch, die beiden Sauerstoffatome ununterscheidbar. Weder die Doppelbindung noch die negative Ladung können demnach im Molekül lokalisiert sein, wie es durch die Strichformeln (↑ Abb. 5.1.2-8a, b) ausgedrückt wird. Die tatsächliche Elektronenverteilung im Molekül liegt zwischen den Elektronenverteilungen der *mesomeren Grenzstrukturen a* und *b*, was für die bindenden

Elektronenpaare durch Formel *c* symbolisiert wird. Die Grenzstrukturen sind keine real existierenden Zustände des Moleküls, sondern Hilfsmittel zur Veranschaulichung der wahren Bindungsverhältnisse.

Im Rahmen einer auf PAULING zurückgehenden quantenmechanischen Näherung, der *v*alence *b*ond- oder *VB-Theorie*, wird die Wellenfunktion eines Moleküls aus einer Linearkombination von *Valenzstrukturen*, d. h. von Wellenfunktionen der mesomeren Grenzstrukturen, erhalten.

Im Unterschied zur Molekülorbitalbehandlung bleibt dabei die qualitative Vorstellung, daß die Bindung durch zwischen bestimmten Kernpaaren lokalisierte Elektronenpaare bewirkt wird, erhalten. Aus diesem Grunde ist für die VB-Methode auch die Bezeichnung *Elektronenpaarmethode* gebräuchlich.

Die Bindungsverhältnisse im Benzen lassen sich vereinfacht schon durch Angabe zweier mesomerer Grenzstrukturen, der Kekuléformeln, beschreiben (Abb. 5.1.2-9a und b).

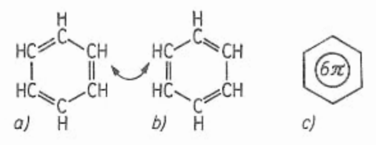


Abb. 5.1.2-9 Mesomerie am Beispiel des Benzens

Der Delokalisierung der 6π-Elektronen über das gesamte Molekül wird durch die Schreibweise entsprechend der Abb. 5.1.2-9c Rechnung getragen. Benzen ist aufgrund der Konjugation seiner π-Elektronen um etwa 88 kJ mol^{-1} energieärmer als das den mesomeren Grenzstrukturen a, b entsprechende hypothetische Cyclohexatrien. Analoge Verhältnisse treten nach HÜCKEL bei allen zyklisch konjugierten Systemen auf, die $(4n + 2)\pi$-Elektronen besitzen ($n = 0, 1, 2, 3, \ldots$). Zyklisch konjugierte Systeme, die diese *Hückelregel* erfüllen, werden *aromatisch* genannt.

5.1.2.6. Delokalisierte π-Bindungen und Hückel-Molekülorbital-Methode. Für ein Molekül mit N Kernen (Kernladungszahlen z_j bzw. z_l) und n Elektronen hat der Hamiltonoperator die Form

$$(7)\quad \hat{H} = -\frac{h^2}{8\pi^2 m}\sum_{i=1}^{n}\Delta_i - \sum_{i=1}^{n}\sum_{j=1}^{N}\frac{z_l e^2}{4\pi\varepsilon_0 r_{ij}} + \sum_{i=1}^{n}\sum_{k=i+1}^{n}\frac{e^2}{4\pi\varepsilon_0 r_{ik}} + \sum_{j=1}^{N}\sum_{l=j+1}^{N}\frac{z_j z_l e^2}{4\pi\varepsilon_0 r_{jl}}.$$

Das erste Glied erfaßt die kinetische Energie der Elektronen, das zweite die Elektronen-Kern-Wechselwirkung, das dritte die Wechselwirkung der Elektronen untereinander und das vierte die Kern-Kern-Wechselwirkung. Selbst für Ethylen C_2H_4, den einfachsten Kohlenwasserstoff mit delokalisierten π-Elektronen, würde eine Näherungsberechnung der erlaubten Energien einen beträchtlichen Rechenaufwand erfordern. Eine Vereinfachung der Rechnung für Moleküle mit konjugierten π-Bindungen läßt sich unter der Annahme, daß die π-Elektronen unabhängig von den σ-Elektronen betrachtet werden können, erzielen. Im Rahmen dieser *π-Elektronen-Näherung* wird der σ-Rumpf als starr angesehen. Für die Bildung des π-Bindungssystems stellt jedes C-Atom ein π-Orbital zur Verfügung. Die recht drastisch erscheinende Näherung, die nicht einmal alle Valenzelektronen der beteiligten Atome berücksichtigt, wird dennoch durch die Tatsache gerechtfertigt, daß sich charakteristische Eigenschaften vieler, insbesondere zyklisch konjugierter planarer Moleküle mit Hilfe des π-Elektronensystems verstehen lassen.

Zu einer weiteren Vereinfachung gelangt man nach HÜCKEL, wenn man im Hamiltonoperator nur Einelektronenoperatoren zuläßt, d. h. die Wechselwirkung der Elektronen untereinander nicht explizit berücksichtigt. Im Rahmen der LCAO-MO-Methode (↑ 5.1.2.1) resultieren Einelektronen-Molekülorbitale als Linearkombination von Atomorbitalen der am π-System beteiligten Atome: $\psi_i = \sum_{i=1}^{n} C_{ij}\Phi_j$. Die Anwendung des Variationsverfahrens (↑ 5.1.2.1.) führt auf eine Säkulardeterminante der Ordnung $n \times n$. Die Lösung liefert die Werte für die Energien des Grundzustands und der angeregten Einelektronenzustände. Im Rahmen der Hückelschen Molekülorbitalmethode (HMO) gelten folgende Annahmen: $H_{ii} = \alpha$; $H_{ij} = \beta_{ij} = \beta$, wenn die Atome i und j benachbart sind; $H_{ij} = 0$, wenn die Atome i und j nicht benachbart sind: $S_{ij} = 1$, wenn $i = j$ und $S_{ij} = 0$, wenn $i \neq j$. Die Integrale α und β sind negative Größen. Sie werden im Rahmen der HMO-Methode nicht berechnet, sondern durch *empirische Parameter* angenähert. α-Werte können z. B. den Ionisierungsenergien für die Elektronen in den entsprechenden Atomorbitalen gleichgesetzt werden.

Für Ethylen hat die Säkulardeterminante die Gestalt $\begin{vmatrix}\alpha - \varepsilon & \beta \\ \beta & \alpha - \varepsilon\end{vmatrix} = 0$; die Eigenwerte ergeben sich zu $\varepsilon_1 = \alpha + \beta$ und $\varepsilon_2 = \alpha - \beta$. Das Einsetzen der Energiewerte in die Säkulargleichungen unter Berücksichtigung der Normierungsbedingung liefert die Koeffizienten $c_1 = c_2 = 1/\sqrt{2}$ und $c_1 = -c_2 = 1/\sqrt{2}$. Es resultieren die Wellenfunktionen $\psi_1 = (1/\sqrt{2})(\Phi_1 + \Phi_2)$ und $\psi_2 = (1/\sqrt{2})(\Phi_1 - \Phi_2)$. Das Energieniveaudiagramm ist in Abb. 5.1.2-10 dargestellt. Das Niveau mit der Energie ε_1 ist durch zwei

Elektronen mit antiparallelem Spin besetzt. α gibt die Energie eines $2p$-Elektrons in einem isolierten C-Atom an.

ε $-\beta$ —— ε_2
α
β —↑↓— ε_1

Abb. 5.1.2-10 HMO-Schema für Ethylen (Erläuterung im Text)

Die HMO-Methode ist ein *halbempirisches Verfahren* zur Molekülberechnung. Vollständige Molekülberechnungen ohne empirische Parameter erfordern einen ungleich höheren Rechenaufwand, da sich die Molekülorbitale um so genauer approximieren lassen, je vollständiger der verwendete Satz von Atomorbitalen ist. Mit Vergrößerung der atomaren Basis vergrößert sich aber auch die Zahl der zu berechnenden Integrale. Bei einem Basissatz von 100 Atomorbitalen sind z. B. zur Berücksichtigung der Elektronenwechselwirkung mehr als 10^7 Zweielektronenintegrale zu lösen. In praktischer Hinsicht haben solche *ab initio-Molekülberechnungen* hauptsächlich den Sinn, Aussagen für Substanzen zu liefern, die experimentell nicht oder nur schwer zugänglich sind.

5.1.3. Polare Atombindungen

5.1.3.1. Übergänge zwischen kovalenter Bindung und Ionenbeziehung. Eine Zuordnung von Verbindungen zu verschiedenen Bindungstypen ist nur in relativ wenigen Fällen zweifelsfrei möglich. So läßt sich z. B. in der Verbindungsreihe FF, SF_2, PF_3, SiF_4, AlF_3, MgF_2, NaF die Bindung im F_2 dem kovalenten Bindungstyp zuordnen, während die Bindung im NaF weitgehend ionischen Charakter hat. In den anderen Stoffen liegen polare Atombindungen mit zunehmend ionischem Anteil vor.

Die Molekülorbitalbehandlung zweiatomiger Moleküle mit verschiedenen Kernen unterscheidet sich nicht grundsätzlich von der Behandlung homonuklearer Moleküle. Die Molekülbahnfunktionen werden ebenfalls durch Linearkombination von Eigenfunktionen der beiden Atome gebildet: $\psi = c_1\Phi_1 \pm c_2\Phi_2$, allerdings sind die Koeffizienten c_i im Falle ungleicher Bindungspartner verschieden. Die resultierenden Molekülorbitale sind nicht mehr symmetrisch bezüglich einer Ebene, die die Kernverbindungslinie halbiert und auf ihr senkrecht steht. Abb. 5.1.3-1 veranschaulicht den Sachverhalt am Beispiel der berechneten Elektronendichteverteilung im H_2- und im LiF-Molekül.

5.1.3.2. Bindungsenergien. Die Enthalpieänderung $\Delta_D H$ bei der Reaktion eines Mols AB gemäß $AB(g) \rightarrow A(g) + B(g)$ wird als molare *Dissoziationsenthalpie* bezeichnet. Sie ist in diesem Fall mit der Bindungsenergie E_{A-B} identisch. Bindungsenergien sind somit zwar experimentell zugänglich, sie müssen aber vor allem bei vielatomigen Molekülen sorgfältig definiert werden, um wirkliche Bedeutung zu haben. Infolge der Äquivalenz aller Bindungen im Methan CH_4 müssen alle C—H-Bindungsenergien E_{C-H} gleich sein: $CH_4\,(g) \rightarrow C\,(g) + 4H\,(g)$; $\Delta_D H = 4E_{C-H}$.

Die *mittlere molare Bindungsenergie* E_{C-H} ist jedoch nicht gleich der Enthalpieänderung bei der Reaktion $CH_4\,(g) \rightarrow CH_3\,(g) + H\,(g)$, da nach Aufbrechen einer Bindung die Natur der drei verbleibenden Bindungen verändert ist. Bindungsenergien sind dennoch von Wert, da sie im Unterschied zu den Dissoziationsenergien einzelner Bindungen oft leicht experimentell, z. B. aus *Verbrennungsenthalpien*, zugänglich sind. Außerdem ergibt sich oft die Gesamtbindungsenergie von Molekülen in guter Näherung additiv aus den mittleren Bindungsenergien der einzelnen Bindungen. Unter dieser Annahme läßt sich z. B. aus der Enthalpieänderung $\Delta_R H$ bei der Reaktion $C_2H_6\,(g) \rightarrow 2C\,(g) + 6H\,(g)$ über $\Delta_R H = 6E_{C-H} + E_{C-C}$ bei Kenntnis von

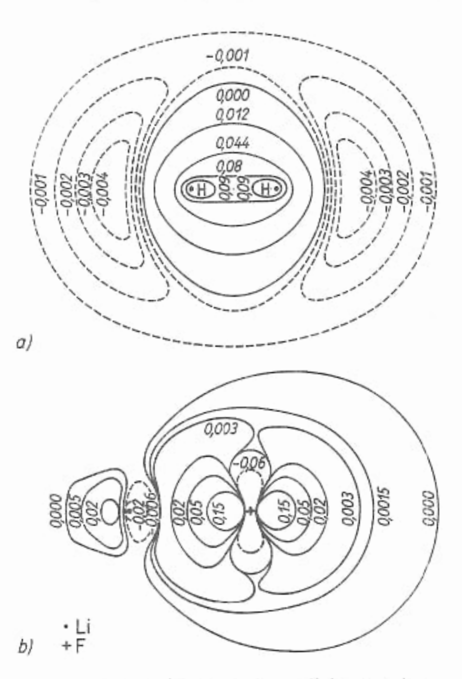


Abb. 5.1.3-1 Differenzladungsdichteverteilung a) für H_2, b) für LiF; *durchgezogene Linien* Elektronendichteerhöhung, *gestrichelte Linien* Elektronendichteerniedrigung gegenüber freien Atomen

E_{C-H} die mittlere molare Bindungsenergie E_{C-C} für C—C-Einfachbindungen errechnen. Eine Vielzahl analoger Rechnungen zeigt (↑ Tab. 7.1.2-3), daß die *Energien von Mehrfachbindungen* höher sind als die von Einfachbindungen, und daß die Stärke einer Bindung mit zunehmender Polarität anwächst.

5.1.3.3. Elektronegativität der Atome und Ionencharakter von Bindungen. Bei Verbindungen AB mit vorwiegend kovalenter Bindung stimmt die Bindungsenergie recht gut mit dem arithmetischen bzw. geometrischen Mittel der Bindungsenergien von AA und BB überein: $E_{A-B} \approx (1/2)(E_{A-A} + E_{B-B}) \approx \sqrt{E_{A-A} \cdot E_{B-B}}$. Die Bindungsenergie polarer Moleküle ist meist größer. Die Größe der Abweichung Δ_{AB} kann als Maß für die Beteiligung ionischer Valenzstrukturen angesehen werden, z. B. gemäß:

$$(8) \quad E_{A-B} = (1/2)(E_{A-A} + E_{B-B}) + \Delta_{AB}.$$

So beträgt die Bindungsenergie der polaren HF-Moleküle 560 kJ mol^{-1}, von H_2 432 kJ mol^{-1} und von F_2 151 kJ mol^{-1}. Sowohl arithmetisches wie geometrisches Mittel liegen mit 292 bzw. 256 kJ mol^{-1} bedeutend niedriger als die experimentell für HF ermittelte Bindungsenergie. Die Differenz Δ_{HF} wird durch den partiellen Ionencharakter der Bindung hervorgerufen. Wie von PAULING empirisch ermittelt wurde, läßt sie sich durch Differenzen zweier für A und B charakteristischer Größen, der *Elektronegativitäten X*, ausdrücken:

$$(9) \quad E_{A-B} = 0{,}5(E_{A-A} + E_{B-B}) + 100\ \text{kJmol}^{-1}(X_A - X_B)^2 - 6{,}5\ \text{kJmol}^{-1}(X_A - X_B)^4.$$

Die Elektronegativität X eines Atoms A kennzeichnet die Kraft, mit der es über eine kovalente Bindung Elektronen vom Bindungspartner anzieht; sie wird als dimensionslose Zahl angegeben.

Da aus 5.1.(8) und (9) nur Elektronegativitätsdifferenzen resultieren, wird ein Zahlenwert (X) willkürlich festgelegt. In der Paulingskala wird dem elektronegativsten Element Fluor ein Wert

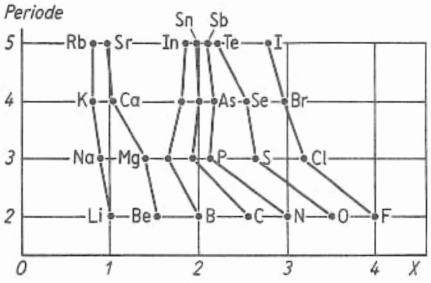


Abb. 5.1.3-2 Elektronegativitäten X ausgewählter Hauptgruppenelemente nach PAULING

$X_F = 4{,}0$ zugeordnet. Für Wasserstoff ergibt sich dann $X_H = 2{,}1$; weitere Werte sind aus Abb. 5.1.3-2 ersichtlich. Man erkennt, daß die Elektronegativität eine periodische Eigenschaft der Elemente ist.

Durch MULLIKEN erhielten die Elektronegativitätswerte eine gewisse theoretische Fundierung; er zeigte, daß die Summen aus dem ersten Ionisierungspotential und der Elektronenaffinität der Atome den Elektronegativitätswerten nach PAULING in recht guter Näherung proportional sind. Vereinzelt auftretende größere Abweichungen waren Anlaß zur Schaffung zahlreicher weiterer Elektronegativitätsskalen. Allen gemeinsam ist, daß der ionische Anteil einer Bindung mit wachsender X-Differenz größer wird, die jeweils für die Ionizität von Bindungen erhaltenen Zahlenwerte unterscheiden sich aber erheblich. Nach dem klassischen Paulingschen Ansatz entsprechen Elektronegativitätsdifferenzen von 1,7 bzw. 0,6 und 3,0 ungefähr 50, 10 und 90% ionischem Charakter der jeweiligen polaren Bindung.

5.1.3.4. Bindungslängen und Teilchenradien. Die Tatsache, daß den Elektronen der Atome und Ionen keine scharf begrenzten Aufenthaltsräume zugeordnet werden können, führt zu Schwierigkeiten bei der Festlegung der Größe von Atomen bzw. Ionen. Genau meßbar sind nur Kernabstände in Molekülen oder Kristallgittern. Handelt es sich dabei um Abstände zwischen gleichen Kernen, ist es üblich, den halben Kernabstand als Teilchenradius zu definieren. Damit wird eine Kugelgestalt der Teilchen impliziert, die für gebundene Atome (s. a. Abb. 5.1.3-1) meist nicht mehr gegeben ist, allenfalls noch bei hohen Koordinationszahlen. Liegen unterschiedliche Atomsorten vor, besteht in der Aufteilung des Abstands auf die Bindungspartner eine erhebliche Willkür, da zur Ableitung z. B. von *Ionenradien* aus den Kernabständen in Ionengittern mindestens ein Ionenradius bekannt sein müßte, aus Strukturuntersuchungen aber stets nur Abstände und daraus Radiensummen oder -differenzen zugänglich sind. Die Vielzahl vorgeschlagener Radiensysteme (↑ 6.4.4.2.), die auf einer Additivität der Teilchenradien zu den experimentell ermittelten Teilchenabständen aufbauen, erlaubt also letztlich keine weitergehenden Aussagen, als sie aus den Abständen selbst zu treffen sind.

Der nichtkugeligen Gestalt gebundener Atome kann man dadurch Rechnung tragen, daß man statt von Radien von Teilchenabmessungen in Bindungsrichtung spricht, damit aber de facto wieder auf die sichere Basis experimentell bestimmter Teilchenabstände zurückgeht.

5.1.3.5. Dipolmomente von Molekülen. Die durch eine unsymmetrische Verteilung der Elektronendichte hervorgerufene Polarität von Bin-

dungen spiegelt sich in deren Dipolmoment wider. Als Vektorsumme der einzelnen Bindungsdipole ergibt sich das Dipolmoment eines Moleküls. CCl_4 ist z. B. ein unpolares Molekül, da sich die vier C—Cl-Bindungsdipole infolge der Tetraedersymmetrie des Moleküls vektoriell gerade zu Null addieren. Für die Ausbildung eines Dipolmoments ist nicht nur die Elektronegativitätsdifferenz der verbundenen Atome verantwortlich. Bei der Überlappung von Orbitalen verschiedener Atome mit unterschiedlicher Größe liegt das Überlappungsgebiet und damit der Ladungsschwerpunkt der bindenden Elektronen außerhalb der Mittelebene auf der Seite des kleineren Atoms. Im HCl-Molekül ist beispielsweise ein derartiger *homöopolarer Dipol* dem durch die Elektronegativitätsdifferenz hervorgerufenen Dipolmoment entgegengerichtet.

5.1.4. Zwischenmolekulare Wechselwirkungen

Auch zwischen Teilchen mit abgesättigten Valenzen bestehen noch Wechselwirkungen. Ihr Vorhandensein zeigt sich z. B. in der Möglichkeit zur Verflüssigung bzw. Erstarrung der Edelgase und aller molekularen Verbindungen bei entsprechender Abkühlung. Die Existenz einer zwischen allen Molekülen und Atomen wirkenden Anziehungskraft wurde bereits 1873 von VAN DER WAALS postuliert und in einer nach ihm benannten Zustandsgleichung für reale Gase berücksichtigt (↑ 6.2.2.). Die im Vergleich zu den Schmelztemperaturen von Ionenkristallen und Metallen relativ niedrige Schmelztemperatur und hohe Flüchtigkeit molekularer Verbindungen weist auf die niedrigen Beträge dieser Wechselwirkungsenergien hin.

5.1.4.1. Bindungskräfte vom van-der-Waals-Typ oder unspezifische Wechselwirkungen. Hierunter faßt man, besonders im chemischen Schrifttum, drei Arten von Wechselwirkungen zwischen dipolbehafteten bzw. dipolfreien Molekülen zusammen: Nach KEESOM wird die mittlere Wechselwirkungsenergie zwischen zwei gleichen Teilchen mit einem permanenten Dipolmoment μ im Abstand r nach $E_K = -(2/3)(\mu^4/r^6)\,1/k_B T$ errechnet, wobei k_B die Boltzmannkonstante und T die absolute Temperatur bedeuten. Die Wechselwirkung führt zu einer bevorzugten gegenseitigen Orientierung der Teilchen, der wiederum die Wärmebewegung entgegenwirkt. Moleküle mit permanenten Dipolmomenten erzeugen in benachbarten Molekülen induzierte Dipolmomente, woraus ebenfalls eine Anziehung resultiert. Die damit verbundene Wechselwirkungsenergie ist bei richtungsunabhängiger Polarisierbarkeit α unabhängig von der Temperatur und beträgt nach DEBYE für zwei gleiche Moleküle $E_D = -2\alpha\mu^2/r^6$.
Verantwortlich für die Anziehung dipolloser Teilchen ist die *Dispersionsenergie*. Sie folgt aus quantenmechanischen Berechnungen, wie sie erstmals von LONDON für die Wechselwirkung zweier H-Atome im großen Abstand durchgeführt wurden. Ein qualitatives Verständnis für die Dispersionsenergie ergibt sich aus der Tatsache, daß selbst Atome mit kugelsymmetrischer Ladungsverteilung oszillierende, also momentane Dipole sind. Diese führen zu elektrischen Feldern, welche in benachbarten Teilchen induzierte Dipole in Phase erzeugen. Es resultiert eine Anziehung der Teilchen. Für zwei unterschiedliche Partikel mit den Polarisierbarkeiten α_1 und α_2 sowie den Ionisierungsenergien I_1 und I_2 errechnet sich die Dispersionsenergie nach $E_L = -(3/2)\cdot(\alpha_1\alpha_2/r^6)\,I_1I_2/(I_1+I_2)$. Die Londonschen Dispersionskräfte sind immer vorhanden und führen stets zu einer Anziehung. Ihr Beitrag zur Bindungsenergie kovalenter Moleküle bzw. zur Gitterenergie von Ionenkristallen ist allerdings gering. Bei Edelgasen sind die schwachen Dispersionskräfte die einzige Art der Wechselwirkung.

5.1.4.2. Elektronen-Donator-Akzeptor-Komplexe. Werden zu einer Lösung von Iod in CCl_4 zunehmende Mengen Alkohol gegeben, so erfolgt eine Farbänderung der Lösung von violett nach braun, entsprechend der kurzwelligen Verschiebung der Absorptionsbande des Iods durch eine spezifische Wechselwirkung des *Elektronenpaardonators* Alkohol mit dem *Elektronenpaarakzeptor* I_2. In der Lösung bildet sich ein *Elektronen-Donator-Akzeptor-Komplex* oder *EDA-Komplex*. Im vorliegenden Fall handelt es sich um einen schwachen EDA-Komplex, der in Lösung im Gleichgewicht mit seinen Komponenten vorliegt. Mit zunehmender Stärke des Komplexes wird der Akzeptor immer mehr an der Elektronenladung des Donators anteilig, was im Grenzfall, z. B. bei den NH_3-Additionsprodukten der Borhalogenide, zur Elektronenstruktur $D^+—A^-$ mit einer kovalenten Bindung führt. Bei der Bildung schwacher EDA-Komplexe treten nur relativ geringe Änderungen der Elektronen- und Schwingungsspektren sowie der NMR-Spektren gegenüber den reinen Komponenten auf. Einige Eigenschaften sind dennoch für den Komplex insgesamt spezifisch, z. B. die Ladungsübertragungs-, charge-transfer- oder *CT-Banden* sowie das Dipolmoment des Molekülkomplexes. Die CT-Bande liegt oft im sichtbaren Bereich des Spektrums. Beispielsweise ist die beim Vereinigen farbloser CCl_4-Lösungen von Anthracen und Tetracyanoethylen $(NC)_2C{=}C(CN)_2$ auftretende Blaufärbung auf eine CT-Bande mit einem Absorptionsmaximum bei 13500 cm^{-1}, entsprechend 740 nm, zurückzuführen. Eine theoretische Begründung der CT-Banden ist auf der Basis der *Mullikenschen VB-Behandlung schwacher EDA-Komplexe* möglich. Es ergibt sich die Eigenfunktion

des Grundzustands ψ_N durch Berücksichtigung zweier Valenzstrukturen: $\psi_N = a\psi_0(D, A) + b\psi_1(D^+ - A^-)$. $\psi_0(DA)$ entspricht der Anordnung der Komponenten in der Gleichgewichtsgeometrie des Komplexes unter der Annahme, daß nur Wechselwirkungen vom van-der-Waals-Typ den Zusammenhalt bewirken. $\psi_1(D^+ - A^-)$ entspricht einem Komplex, in dem ein Elektron völlig vom Donator zum Akzeptor übergegangen ist. Entgegen früheren Auffassungen gilt es heute aufgrund experimenteller und quantenchemischer Untersuchungen als erwiesen, daß die Eigenschaften des Grundzustands schwacher EDA-Komplexe in starkem Maße durch die Nichtbindungs-Eigenfunktion $\psi_0(D, A)$ bestimmt werden. Im Unterschied dazu werden die Eigenschaften des angeregten Zustands vorwiegend durch die Eigenfunktion $\psi_1(D^+ - A^-)$ bestimmt. Danach resultiert die CT-Bande im wesentlichen aus einem intermolekularen Übergang eines Elektrons vom Donator zum Akzeptor. Damit wird unmittelbar verständlich, daß die Energie der CT-Absorption vom Ionisierungspotential des Donators und der Elektronenaffinität des Akzeptors abhängig ist.

5.1.4.3. Wasserstoffbrückenbindungen. Beim Auftragen der Siedetemperaturen oder auch der Verdampfungsenthalpien der Wasserstoffverbindungen der Elemente der 4. bis 7. Hauptgruppe des PSE gegen das Molekulargewicht fällt auf, daß die Werte für H_2O, HF und NH_3 aus der Reihe fallen (Abb. 5.1.4-1). Bei diesen Verbindungen wirkt eine zusätzliche intermolekulare Anziehung besonders stark, die als Wasserstoffbrückenbindung bezeichnet wird.

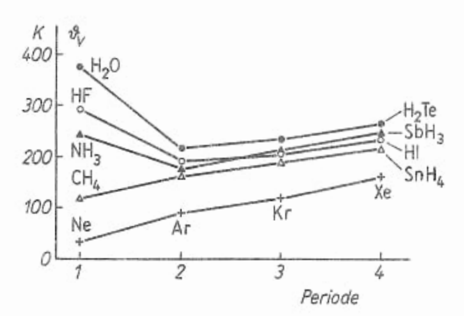


Abb. 5.1.4-1 Siedetemperaturen ϑ_V von Wasserstoffverbindungen und (zum Vergleich) von Edelgasen

Derartige Anziehungskräfte treten immer dann auf, wenn Wasserstoff an ein stark elektronegatives Element X gebunden ist und die Möglichkeit zur Wechselwirkung mit einem zweiten elektronegativen Element Y besteht: X–H...Y.

Die Energie von Wasserstoffbrückenbindungen beträgt meist 20 bis 40 kJmol^{-1}.

In Abb. 5.1.4-2 sind einige Beispiele für Wasserstoffbrückenbindungen angegeben. Intermolekulare Wasserstoffbrücken im Eis erstrecken sich über viele Moleküle. Jedes O-Atom im Eis ist tetraedrisch von vier H-Atomen umgeben. Erstrecken sich intermolekulare

Abb. 5.1.4-2 Wasserstoffbrückenbindung; intermolekular a) bei Eis b) bei Essigsäure; intramolekular c) bei Salicylsäure

H-Brücken nur über zwei Moleküle, kommt es zur Ausbildung von Dimeren. Befinden sich die Atome X und Y innerhalb eines Moleküls in geeigneter sterischer Anordnung, werden intramolekulare H-Brücken ausgebildet. Eine intramolekulare H-Brücke ist z. B. verantwortlich für die im Vergleich zur Löslichkeit des *m*- und *p*-Isomeren geringe Wasserlöslichkeit der *o*-Hydroxy-Benzoesäure oder Salicylsäure. Durch die H-Brücke erscheint die Salicylsäure trotz des größeren Gesamtmoments weniger polar als ihre Isomeren. Im Unterschied zu intermolekularen H-Brücken bleiben intramolekulare Brücken beim Abnehmen der Konzentration der brückenbildenden Substanz in Lösung erhalten.

Extrem starke H-Brücken mit etwa 200 kJ mol^{-1} Bindungsenergie treten im $H_9O_4^+$-Ion, dem hydratisierten Proton, und im Anion des Salzes KHF_2 auf:

[F—H----F]$^\ominus$ [F----H—F]$^\ominus$

Im HF_2^--Ion ist die H-Brücke symmetrisch. Der Abstand zwischen den beiden F-Atomen ist mit 226 pm extrem kurz. Normale X...H...Y-Brücken haben Bindungslängen zwischen 275 und 300 pm. Die Erforschung der Natur der

Wasserstoffbrückenbindung hat in den letzten beiden Jahrzehnten große Fortschritte gemacht, da H-Brücken bei Aufbau und Umwandlung biologischer Strukturen eine große Rolle spielen, z. B. bei den Nukleinsäuren als den Trägern der genetischen Information. Die H-Brücken zwischen den komplementären Basenpaaren der Nukleinsäuren sind einerseits stark genug, um die Doppelhelixstruktur zweier Nukleinsäurefäden zu stabilisieren, andererseits sind sie schwach genug, um bei der Replikation eine Trennung der Doppelstränge zu ermöglichen.

5.2. Struktur der Moleküle

5.2.1. Symmetrie von Molekülen

Trotz der riesigen Zahl chemischer Verbindungen unterschiedlichster Zusammensetzung und Struktur ist über die Symmetrie der jeweiligen Baugruppen eine Klassifizierung möglich.

Unter Symmetrie eines Moleküls versteht man das Vorliegen struktureller Regelmäßigkeiten, die eine Vertauschung bestimmter Atome oder Atomgruppen mit anderen ohne Änderung des Aussehens des Moleküls zulassen.

5.2.1.1. Symmetrieoperationen und Symmetrieelemente. Eine Symmetrieoperation ist eine geometrische Operation, durch die ein Körper in eine zur Ausgangslage äquivalente oder mit ihr identische Lage überführt wird. Als Grundtypen von Symmetrieoperationen sind für Moleküle die Drehung um eine Achse und die Spiegelung an einer Ebene zu nennen, bei kristallinen Festkörpern (↑ 6.4.1.4.) zusätzlich die Translation um einen bestimmten Betrag. Die geometrischen Orte, auf die sich die Symmetrieoperationen beziehen, also Ebenen, Achsen oder Punkte, werden Symmetrieelemente genannt. Da bei allen Symmetrieoperationen, die für Moleküle denkbar sind, wenigstens ein Punkt des Moleküls ortsfest bleibt, hebt man die für Moleküle wichtigen Symmetrieelemente durch die Bezeichnung Punktsymmetrieelemente aus der Gesamtheit heraus.

5.2.1.1.1. *Rotation um eine Drehachse.* Ein Molekül enthält eine n-zählige Drehachse C_n, wenn bei Drehung um eine durch das Molekül gelegte Achse um den Winkel $360°/n$ das Molekül in eine von der Ausgangslage ununterscheidbare Lage übergeht. Das trigonal-ebene Bortrifluorid enthält z. B. eine dreizählige Achse C_3, die senkrecht zur Molekülebene orientiert ist und durch das zentrale Boratom geht (Abb. 5.2.1-1). Das Symbol C_3 steht dabei nicht nur für das Symmetrieelement, sondern auch für die zugehörige Symmetrieoperation, die Rotation um $360°/3$ um die dreizählige Achse. Für die Rotation um $2 \times 360°/3$ steht dementsprechend das Symbol $C_3C_3 = C_3^2$, für $3 \times 360°/3$ schreibt man C_3^3. Die Operation C_3^3 erzeugt die mit der Ausgangslage des Moleküls identische Lage; sie ist damit ein Sonderfall der Identitätsoperation E. Allgemein gilt: $C_n^n = E$.

5.2.1.1.2. *Spiegelung an einer Spiegelebene.* Eine Spiegelebene σ teilt ein Molekül in zwei spiegelbildliche Hälften. Damit ist die Forderung verbunden, daß die Zahl der gleichartigen Atome außerhalb der Spiegelebene nur geradzahlig sein kann, denn für jedes Atom muß ein gleichartiges auf der anderen Seite der Ebene existieren. Einzelatome müssen in der Spiegelebene liegen, da sie nur dann bei Ausführung der Spiegelung ihre Lage nicht ändern. Lineare Moleküle wie CO_2 haben unendlich viele Spiegelebenen, die alle die Molekülachse enthalten. Das trigonal ebene BF_3 hat außer einer Spiegelebene, die alle Atome enthält, drei weitere Spiegelebenen senkrecht zur Molekülebene und parallel zur dreizähligen Achse des Moleküls, die das zentrale B-Atom und jeweils ein F-Atom enthalten und den Winkel zwischen den beiden anderen F-Atomen halbieren. Allgemein werden Spiegelebenen, die parallel zur Haupt- oder Referenzachse, d. h. zur Achse mit der höchsten Zähligkeit n verlaufen, als vertikale Spiegelebenen σ_v bezeichnet. Spiegelebenen senkrecht zur Hauptachse des Moleküls sind horizontale Spiegelebenen σ_h. Da die dreizählige Achse im BF_3 die Hauptachse ist, hat das Molekül drei σ_v und eine σ_h.

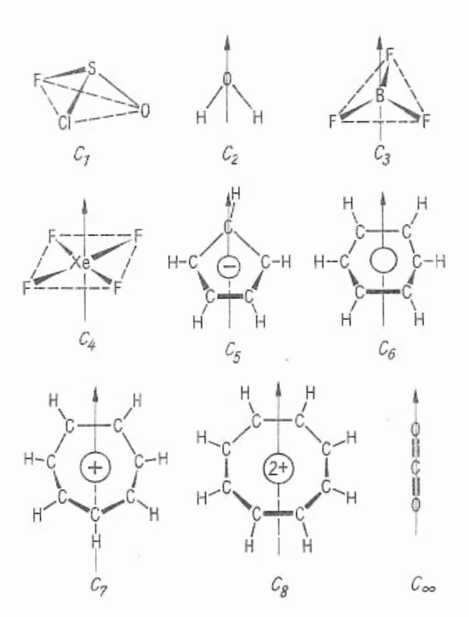


Abb. 5.2.1-1 Moleküle mit Drehachsen unterschiedlicher Zähligkeit

5.2.1.1.3. *Drehspiegelung und Drehspiegelachsen.* Eine Drehspiegelung erfolgt in zwei Schritten: Drehung um eine Achse und Spiegelung an einer Ebene senkrecht zur Drehachse. Das Symbol für die Operation und das zugehörige Symmetrieelement ist S_n, wobei n wieder die Zähligkeit der Achse angibt. Moleküle, die eine Drehachse C_n und unabhängig von dieser eine senkrecht dazu orientierte Spiegelebene aufweisen, haben immer auch eine Drehspiegelachse S_n. Dies ist allerdings keine notwendige Bedingung; z. B. enthält das tetraedrische CCl_4-Molekül eine S_4, obwohl weder eine C_4 noch eine senkrecht dazu orientierte Spiegelebene σ vorliegt. Ein wichtiger Spezialfall ist die zweizählige Drehspiegelachse $S_2 = C_2\sigma_h = \sigma_h C_2$; sie ist gleichbedeutend mit der Inversion i.

5.2.1.1.4. *Inversion i an einem Symmetriezentrum.* Durch die Inversion wird jeder Atomschwerpunkt am Schnittpunkt der Drehachse C_2 mit der Spiegelebene σ_h gespiegelt. Beispiele für Moleküle mit einem Inversionszentrum sind das hexagonal-ebene Benzen, C_6H_6 und das lineare CO_2-Molekül (↑ Abb. 5.2.1-1).

5.2.1.2. Punktgruppen von Molekülen. Die Symmetrie des H_2O-Moleküls (↑ Abb. 5.2.1-1) wird durch den Satz der Symmetrieoperationen E, C_2, σ_v und σ_v' eindeutig charakterisiert. Die Gesamtheit aller Symmetrieoperationen an einem Molekül nennt man dessen *Symmetriegruppe*, wobei die einzelnen Symmetrieoperationen, also nicht die Symmetrieelemente, Elemente dieser Gruppe sind.

Zur Diskussion der Molekülsymmetrie können die mathematischen Methoden der *Gruppentheorie* herangezogen werden. Da die Lage des Molekülschwerpunkts, in dem sich alle Symmetrieelemente schneiden, durch die Symmetrieoperationen nicht verändert wird, werden die Symmetriegruppen der Moleküle auch als *Punktgruppen* bezeichnet. Ihre Nomenklatur geht auf SCHOENFLIES zurück, während zur Kennzeichnung der Symmetrie von Kristallen die von MAUGUIN eingeführten Symbole gebräuchlicher sind (↑ 6.4.1.3.). Jede Punktgruppe erhält ein Symbol, welches in roher Form die Symmetrie der Gruppe beschreibt. Unter Anwendung des in der Abb. 5.2.1-2 angegebenen Algorithmus gelingt es meist sehr leicht, die Punktgruppe eines Moleküls zu ermitteln. Spezielle Punktgruppen werden sehr schnell erkannt: Lineare Moleküle ohne σ_h gehören zur Punktgruppe $C_{\infty h}$, bei zusätzlicher Anwesenheit einer σ_h resultiert die Punktgruppe $D_{\infty h}$. Die kubischen Punktgruppen T_d bzw. O_h fassen die 24 Symmetrieoperationen am regulären Tetraeder bzw. 48 am regulären Oktaeder zusammen. Die Ermittlung der Punktgruppe eines Moleküls soll am Beispiel des Ruthenocens $(C_5H_5)_2Ru$ (s. a. Abb. 5.2.1-1), veranschaulicht werden: Das Molekül ist nicht linear und gehört keiner kubischen Gruppe an. Die Achse höchster Zähligkeit ist eine C_5. Es gibt keine S_{10} parallel zur C_5. Es existieren fünf zweizählige Achsen senkrecht zur C_5, folglich gehört das Molekül zu einer D-Gruppe. Die zusätzliche Existenz einer σ_h führt auf die Punktgruppe D_{5h}.

5.2.1.3. Elektronenpaarabstoßung und Molekülgestalt. Viele Molekülstrukturen lassen sich unter der Annahme verstehen, daß die Molekülgeometrie hauptsächlich durch die Wechselwirkungen zwischen den Elektronenpaaren in der Valenzschale eines Zentralatoms bestimmt wird. Aufbauend auf dieser Vorstellung wurden von GILLESPIE Regeln zur Vorhersage von Molekülstrukturen, vorzugsweise für anorganische Verbindungen der Hauptgruppenelemente, angegeben. Die wichtigste dieser *V*alence *S*hell *E*lektron *P*air *R*epulsion- oder *VSEPR-Regeln* lautet: Die Elektronenpaare in einer Valenzschale ordnen sich in möglichst großer Entfernung voneinander an. Die Rechtfertigung dafür liefert das *Pauliprinzip*. Unter der vereinfachenden Annahme, daß sich die Elektronenpaare in der Valenzschale alle in gleichem mittlerem Abstand vom Kern aufhalten, ergeben sich für

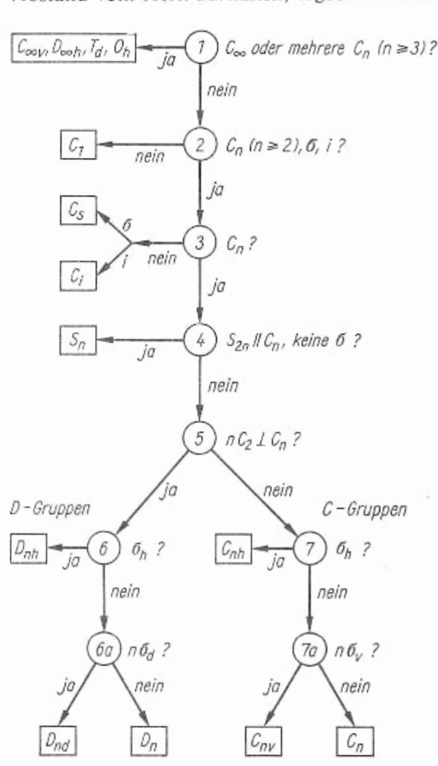


Abb. 5.2.1-2 Algorithmus zur Bestimmung der Punktgruppe eines Moleküls

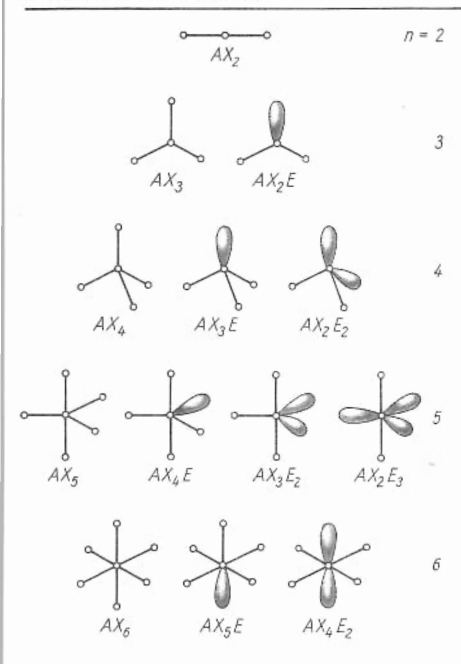


Abb. 5.2.1-3 Bevorzugte Anordnung von *n* Elektronenpaaren und Molekülgestalt

zwei bis sechs Elektronenpaare in jeweils maximaler Entfernung voneinander die in Abb. 5.2.1-3 wiedergegebenen Anordnungen: Sind nicht alle *n* Elektronenpaare an Bindungen beteiligt, resultieren von den Strukturen AX_n abweichende Molekülgestalten. Substanzen gleichen Formeltyps AX_2 treten z. B. je nach Zahl der freien Elektronenpaare E linear ($E = 0$) oder unterschiedlich gewinkelt ($E = 1, 2$) auf.

Für das genauere Verständnis der Strukturen solcher Moleküle AX_mE_{n-m} ist wichtig, daß die nichtbindenden Elektronenpaare E mehr Raum um das Zentralatom beanspruchen als die bindenden. Sie unterliegen nur der Anziehung durch *einen* Kern, während die Aufenthaltswahrscheinlichkeit der bindenden Paare vor allem nahe der jeweiligen Bindungsachse hoch ist. Der Winkel zwischen den verbleibenden Bindungen wird also bei Ersatz bindender durch einsame Elektronenpaare kleiner. Beispielsweise nimmt in der Reihe CH_4, NH_3, H_2O der Bindungswinkel von 109,5° über 107,3° auf 104,5° ab.

In Molekülen mit mehreren möglichen Positionen für freie Elektronenpaare nehmen diese bevorzugt solche Positionen ein, in denen die Wechselwirkung mit anderen Elektronenpaaren möglichst klein ist. Deshalb sind die einsamen Elektronenpaare in AX_4E_2-Molekülen transständig angeordnet; es resultiert eine quadratisch planare Molekülstruktur.

Die Gestalt von Molekülen mit Mehrfachbindungen wird verständlich, wenn man die Elektronenpaare, die die Mehrfachbindungen bilden, zusammenfaßt und wie ein einzelnes Elektronenpaar behandelt, welches allerdings infolge seines größeren Raumbedarfs stärker abstoßend auf benachbarte Bindungen wirkt. Im Phosgen, $Cl_2C{=}O$, beträgt deshalb der Cl—C—Cl-Winkel nur 111,3°, während die Cl—C—O-Winkel auf 124,3° aufgeweitet sind. Moleküle mit fünf Elektronenpaaren in der Valenzschale des Zentralatoms weisen dadurch Besonderheiten auf, daß axiale und äquatoriale Positionen einer trigonalen Bipyramide nicht gleichwertig sind. Die axialen Positionen haben drei nächste Nachbarn im Winkel von 90° gegenüber zwei 90°-Nachbarn bei den äquatorialen Positionen. Die nichtbindenden Elektronenpaare bevorzugen deshalb die äquatorialen Stellungen.

Tab. 5.2.1-4 VSEPR-Regeln und Molekülgestalt

n	Typ	Gestalt	Punktgruppe	Beispiel(e)
2	AX_2	*linear*	$D_{\infty h}$	HgX_2, ZnX_2, $Ag(CN)_2^-$
3	AX_3	*dreieckig*	D_{3h}	BCl_3, BF_3
	AX_2E	V-förmig	C_{2v}	$SnCl_2$, PbI_2
4	AX_4	*tetraedrisch*	T_d	BF_4^-, CH_4, $TiCl_4$
	AX_3E	*trigonal*-pyramidal	C_{3v}	NH_3, PF_3, CH_3^+
	AX_2E_2	V-förmig	C_{2v}	H_2O, H_2S, CCl_2
5	AX_5	*trigonal-bipyramidal*	D_{3h}	PF_5, PCl_5, $TaCl_5$
	AX_4E	irreguläre Tetraeder	C_{2v}	SF_4
	AX_3E_2	T-förmig	C_{2v}	ClF_3, BrF_3
	AX_2E_3	linear	$D_{\infty h}$	ICl_2^-, I_3^-
6	AX_6	*oktaedrisch*	O_h	SF_6, WCl_6, FeF_6^{3-}
	AX_5E	quadratisch-pyramidal	C_{4v}	BrF_5, IF_5
	AX_4E_2	quadratisch-planar	D_{4h}	ICl_4^-, BrF_4^-

In Tab. 5.2.1-4 sind zahlreiche Beispiele für die mit den VSEPR-Regeln übereinstimmenden Molekülgestalten zusammengestellt; die ungefähre Orientierung aller *n* Elektronenpaare zueinander geht dabei jeweils aus dem ersten Begriff der dritten Spalte hervor.

5.2.2. Isomerie bei Molekülverbindungen

Die Summenformeln chemischer Verbindungen sagen im allgemeinen noch nichts über deren inneren Aufbau aus. Um die wesentlichen Merkmale der Molekülstruktur widerspiegeln zu können, sind zusätzlich Angaben zur Konstitution, Konfiguration und Konformation erforderlich.

Die *Konstitution* gibt die Verknüpfung der einzelnen Atome im Molekül an, die *Konfiguration* die bei einer gegebenen Verknüpfung noch möglichen Unterschiede in der stabilen geometrischen Anordnung bestimmter Molekülteile zueinander, und die *Konformation* die bei gegebener Konstitution und Konfiguration bevorzugte Stellung durch Einfachbindungen verknüpfter Molekülteile zueinander.

Isomerie im weitesten Sinne liegt vor, wenn sich Verbindungen gleicher Summenformel in einem oder mehreren der genannten Strukturmerkmale unterscheiden. Sie gibt sich an den unterschiedlichen physikalischen und chemischen Eigenschaften der Isomeren zu erkennen, wobei im allgemeinen die stärkste Differenzierung bei Konstitutions-, die geringste bei Konformationsisomeren auftritt. An ausgewählten Beispielen soll das veranschaulicht werden.

```
    H    H    H    H    H
    |    |    |    |    |
H — C1 — C2 — C3 — C4 — C5 — H
    |    |    |    |    |
    H    H    H    H    H
```
n-Pentan (36°C)

```
         H
         |
      H— C —H
    H    |    H    H
    |    |    |    |
H — C1 — C2 — C3 — C4 — H
    |    |    |    |
    H    H    H    H
```
2-Methyl-Butan, Isopentan (28°C)

```
         H
         |
      H— C —H
    H    |    H
    |    |    |
H — C1 — C2 — C3 — H
    |    |    |
    H    |    H
      H— C —H
         |
         H
```
2,2-Dimethyl-Propan, Neopentan (9,5°C)

Abb. 5.2.2-1 Konstitutionsformeln und Siedetemperaturen der Isomeren des Pentans

5.2.2.1. Konstitutionsisomerie. Butanol-(1) und Diethylether gehören trotz gleicher Summenformel $C_4H_{10}O$ ganz verschiedenen Stoffklassen an, da der Sauerstoff unterschiedlich in das Molekül eingebunden ist, wie aus den Konstitutionsformeln (↑ 5.1.2.5) ersichtlich ist:

```
  H  H  H  H
  |  |  |  |
H—C—C—C—C—OH
  |  |  |  |
  H  H  H  H
```
Butanol-(1) (117,5 °C)

```
  H  H     H  H
  |  |     |  |
H—C—C—O—C—C—H
  |  |     |  |
  H  H     H  H
```
Diethylether (35,5 °C)

```
  H  H  H  H
  |  |  |  |
H—C—C—C—C—H
  |  |  |  |
  H  H OH  H
```
Butanol-(2) (100 °C)

```
  H  H  H     H
  |  |  |     |
H—C—C—C—O—C—H
  |  |  |     |
  H  H  H     H
```
n-Propyl-methyl-ether (39 °C)

Die bei Alkoholen mögliche Assoziation über Wasserstoffbrücken bewirkt die, jeweils in Klammern angegebene, relativ hohe Siedetemperatur des *n*-Butanols. Geringere Unterschiede ergeben sich innerhalb der Paare Butanol-(1) Butanol-(2) bzw. Diethyl- und Propyl-methylether, da hier lediglich die Stellung, *Stellungsisomerie*, nicht die Natur der funktionellen Gruppen verändert ist. Bei Butanol kann die Verzweigung des Alkylrests zu einem weiteren Isomeren, dem Isobutanol oder 2-Methylpropanol-(1), führen. Diese *Skelett-* oder *Kettenisomerie* ist bei den höheren Alkanen besonders ausgeprägt. Abb. 5.2.2-1 zeigt die drei Konstitutionsisomeren des Pentans, C_5H_{12}. Zur Bezeichnung der Verbindungen wird nach der *Genfer Nomenklatur* die längste C-Kette durchnumeriert und jede Verzweigung gesondert gekennzeichnet; daneben sind auch Trivialnamen gebräuchlich. Mit wachsender Zahl der C-Atome steigt die Zahl der Skelettisomeren rasch an, z. B. beim Octan C_8H_{18} auf 18, beim Decan $C_{10}H_{22}$ auf 75.

Bei anorganischen Komplexverbindungen treten verschiedene Formen der Konstitutionsisomerie auf: Verbindungen gleicher Zusammensetzung, die in Lösung in verschiedene

Ionen zerfallen, werden als *Ionisationsisomere* bezeichnet. So tritt bei den Verbindungen $[Co(NH_3)_4Cl_2]NO_2$ und $[Co(NH_3)_4Cl(NO_2)]Cl$ alternativ Nitrit oder Chlorid als Anion auf. Ein spezieller Fall ist die *Hydratisomerie*, wobei Aquogruppen H_2O entweder sämtlich über Sauerstoff an das Kation koordiniert oder teilweise auch relativ locker als Kristallwasser gebunden sind: So ist eine wäßrige Lösung von $[Cr(OH_2)_6]Cl_3$ violett gefärbt. Bei Zusatz von Silberionen werden alle Chloridionen sofort als schwerlösliches AgCl ausgefällt, aus einer Lösung des grünen $[Cr(OH_2)_4Cl_2]Cl \cdot 2\,H_2O$ dagegen nur ein Drittel. Bei Verbindungen mit komplexen Kationen und Anionen kann die Verteilung der Liganden auf die beiden Koordinationssphären variieren. Es treten *Koordinationsisomere* auf, z. B. $[Co(NH_3)_6]\,[Cr(CN)_6]$ und $[Cr(NH_3)_6]\,[Co(CN)_6]$.

Ist ein Ligand in verschiedenen Komplexen gleicher Summenformel auf unterschiedliche Weise am Zentralatom koordiniert, so liegt die sogenannte *Salzisomerie* vor. Eine NO_2^--Gruppe kann z. B. als Nitrogruppe über den Stickstoff, als Nitritogruppe über den Sauerstoff gebunden werden.

Liegen Konstitutionsisomere nebeneinander im chemischen Gleichgewicht vor, spricht man von *Tautomeren.*

Bei der *Keto-Enol-Tautomerie* wandeln sich die Isomeren durch Wanderung eines Protons bei gleichzeitiger Umordnung der Bindungselektronen ineinander um. Während einfache Ketone wie Aceton vorwiegend in der Ketoform vorliegen, kann bei 1,3-Dicarbonylverbindungen wie dem Acetylaceton die Enolform im Gleichgewicht begünstigt sein:

CH_3 CH_2 CH_3 ⇄ CH_3 CH CH_3
C C ⇄ C C
‖ ‖ | ‖
O O O O
H

Ketoform (28 %) Enolform (72 %).

Die Enolform ist hier durch Ausbildung einer intramolekularen Wasserstoffbrücke und konjugierter Doppelbindungen stabilisiert. Die Lage eines Keto-Enol-Gleichgewichts ist lösungsmittel- und temperaturabhängig; die hier angegebenen Zahlenwerte gelten für die reine Flüssigkeit bei Raumtemperatur.

Reversible Umlagerungen von Molekülen, bei denen lediglich eine Umgruppierung von Valenzelektronen erfolgt, werden als Valenzisomerisierungen bezeichnet. Sie führen zu *Valenzisomeren* oder *Valenztautomeren.* Ein Beispiel ist die *Copeumlagerung* (Abb. 5.2.2.-2a). Für R=H sind Reaktant und Produkt identisch. Es liegt eine entartete Umwandlung oder *Topomerisierung* vor. Ein extremes Beispiel für ein Molekül mit sehr schnellen und reversiblen Topomerisierungen ist das *Bullvalen*, $C_{10}H_{10}$. In Abb. 5.2.2-2b ist eine mögliche Umwandlung zwischen zwei Isomeren veranschaulicht.

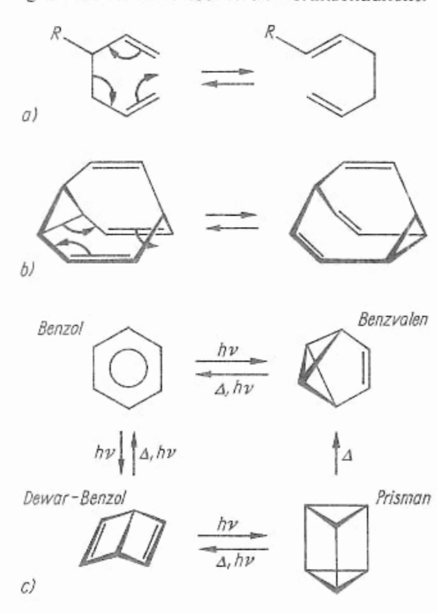


Abb. 5.2.2-2 Valenzisomerie; a) Cope-Umlagerung, b) Topomerisierung bei Bullvalen, c) Valenzisomere des Benzens und ihre fotochemische *hν* bzw. thermische *Δ* Umwandlung

Insgesamt gibt es $10!/3 = 1\,209\,600$ strukturgleiche Isomere des Bullvalens. Jedes C-Atom läßt sich mit jedem anderen durch Cope-Umlagerung beliebig kombinieren. Bei Raumtemperatur verlaufen die Umlagerungen so schnell, daß trotz vier unterschiedlicher Wasserstoffpositionen im ^{1}H-NMR-Spektrum (↑ 4.4.3.) nur ein Signal beobachtet wird.

Interessant sind in diesem Zusammenhang die inzwischen gelungenen Synthesen von Valenzisomeren des Benzens, die teilweise klassischen Strukturvorschlägen für das Benzenmolekül entsprechen. Die Isomeren weisen trotz der hohen Ringspannungen eine relativ hohe thermische Stabilität auf (Abb. 5.2.2-2c).

5.2.2.2. Konfigurationsisomerie. Konfigurationsisomere sind Moleküle, die bei gleicher Konstitution unterschiedliche räumliche Anordnungen realisieren, wobei aber mögliche Rotationen um Einfachbindungen außer Betracht gelassen werden. Wichtige Fälle sind die optische und die geometrische Isomerie, letztere vor allem als cis-trans-Isomerie.

5.2.2.2.1. *Optische Isomerie.* Moleküle, die mit ihrem Spiegelbild nicht kongruent sind, werden als chiral bezeichnet. Sie sind *optisch aktiv*, d. h., sie drehen bei Wechselwirkung mit linear polarisiertem Licht dessen Polarisationsebene um

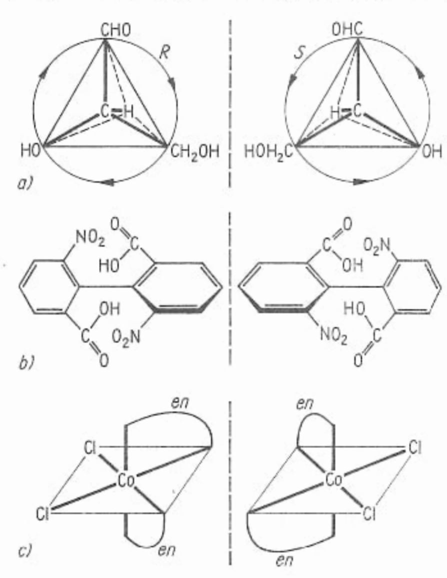


Abb. 5.2.2-3 Enantiomere von a) Glyceraldehyd, b) Dinitrodiphensäure, c) $[Co(en)_2Cl_2]Cl$

einen bestimmten Betrag. *Enantiomere*, d. h., Konfigurationsisomere, die sich wie Bild und Spiegelbild verhalten, drehen die Polarisierungsebene um den gleichen Betrag, aber im entgegengesetzten Drehsinn. Die Drehung wird durch unterschiedliche Brechungsindizes der Enantiomeren für die beiden zirkular polarisierten Komponenten des linear polarisierten Lichtstrahls, die *Zirkulardoppelbrechung*, hervorgerufen. *Chiralität* als notwendige und hinreichende Bedingung für optische Aktivität ist bei Molekülen der Punktgruppen C_1, C_n oder D_n zu erwarten. Derartige Moleküle enthalten keine Spiegelebene, kein Symmetriezentrum und keine Drehspiegelachse. Verbindung *a* in Abb. 5.2.2-3, der Glyceraldehyd, enthält ein asymmetrisches, d. h. mit vier unterschiedlichen Substituenten verbundenes C-Atom. Bei Verbindung *b* ist die freie Drehbarkeit um die zentrale σ-Bindung infolge sterischer Behinderung der Substituenten blockiert. Das Molekül weist darum weder eine Spiegelebene noch ein Inversionszentrum auf. Eine Spaltung in *optische Antipoden* ist möglich. Bei anorganischen Komplexverbindungen, deren Moleküle kein Symmetriezentrum haben, sind ebenfalls optische Isomere zu erwarten. In vielen Fällen gelang eine Trennung in die Isomeren, z. B. beim cis-$[Co(en)_2Cl_2]Cl$, dem cis-Dichloro-bis-ethylendiamin-cobalt(III)-chlorid.

Zur eindeutigen Kennzeichnung der absoluten Konfiguration optisch aktiver Verbindungen wird die von CAHN, INGOLD und PRELOG entwickelte Nomenklatur verwendet: Dazu werden die vier Substituenten nach fallender Priorität geordnet, d. h. nach fallender Ordnungszahl der am Asymmetriezentrum direkt gebundenen Atome. Bei gleichen ersten Atomen werden die Ordnungszahlen der an sie gebundenen zweiten, eventuell noch der dritten Atome verglichen. Sind dabei zwei Atome durch eine Doppel- oder Dreifachbindung verknüpft, werden beide verdoppelt oder verdreifacht, z. B. $—CH{=}CH_2 \mathrel{\hat{=}} —\underset{|\atop C}{C}H—\underset{|\atop C}{C}H_2$. Für Glyceraldehyd (↑ Abb. 5.2.2-3) liefert die *Sequenzregel* die Prioritätsreihe $OH > CHO > CH_2OH > H$. Anschließend wird das Molekül aus einer Richtung betrachtet, die dem Substituent mit der niedrigsten Priorität entgegengesetzt ist. Sind die übrigen Substituenten nach fallender Priorität im Uhrzeigersinn angeordnet, liegt die *R*-Form [*rectus* lat., rechts, richtig] vor, bei entgegengesetzter Anordnung die *S*-Form [*sinister* lat., links, verkehrt].

Bei Verbindungen mit mehreren optisch aktiven Atomen ist das Verfahren auf jedes Chiralitätszentrum anzuwenden.

Konfigurationsisomere, die nicht enantiomer sind, bezeichnet man als *Diastereoisomere*. Sie treten z. B. bei Verbindungen mit mehreren asymmetrischen Atomen auf. So gibt es eine zu den beiden optisch aktiven Weinsäuren diastereoisomere meso-Weinsäure, die selbst nicht optisch aktiv ist. Diastereoisomere unterscheiden sich in ihren physikalischen und chemischen Eigenschaften stärker als Enantiomere. Dies kann man zur Spaltung von *Racemformen*, d. h. Stoffen, die zu gleichen Anteilen aus zwei optischen Antipoden bestehen, ausnutzen. Setzt man z. B. die Racemform einer optisch aktiven Säure mit einer optisch aktiven Base um, so lassen sich die entstehenden diastereoisomeren Salze infolge unterschiedlicher Löslichkeit durch fraktionierte Kristallisation trennen. Anschließende Spaltung der Salze liefert die reinen Enantiomeren der Säure.

5.2.2.2.2. *cis-trans-Isomerie.* Konfigurationsisomere werden als cis-trans-Isomere bezeichnet, wenn sie sich in der Anordnung der Substituenten bezüglich einer spezifischen, durch das Molekül gelegten und als starr anzusehenden Ebene voneinander unterscheiden.

In cis-Stilben z. B. befinden sich beide Phenylringe auf einer Seite einer senkrecht zur Molekülebene orientierten und die C—C-Doppel-

bindung einschließenden Ebene. Es zeigt ein anderes Absorptionsspektrum im ultravioletten Spektralbereich (Abb. 5.2.2-4) als trans-Stilben, in dem die Phenylringe auf entgegengesetzten Seiten dieser Ebene angeordnet sind. Die Umwandlung der einen Form in die andere setzt ein zeitweiliges Lösen der π-Bindung voraus, z. B. als Folge der Absorption von Licht. Die cis-trans-Nomenklatur ist nur dann eindeutig, wenn zwei gleiche Substituenten auftreten, deren relative Lage betrachtet wird. Befinden sich wie im 1-Brom-2-Chlor-1-Iod-Ethylen (Abb. 5.2.2-5) nur ungleiche Substituenten an einer C—C-Doppelbindung, unterliegt die Zuordnung von cis-trans-Isomeren einer gewissen Willkür. Eine eindeutige Bezeichnung ist mit Hilfe der *E-Z*-Nomenklatur möglich. Danach wird nach der Sequenzregel zunächst die Priorität der zwei Substituenten an jedem Atom der Doppelbindung ermittelt. Die Konfiguration mit zwei Substituenten höherer Priorität an entgegengesetzten Seiten der Doppelbindung wird mit *E* bezeichnet. Befinden sich diese beiden Substituenten auf einer Seite der Doppelbindung, d. h. zusammen, liegt das *Z*-Isomere vor. Im 1-Brom-2-Chlor-1-Iod-Ethylen hat am C-Atom 1 Iod und am C-Atom 2 Chlor die höhere Priorität; daraus resultiert die in Abb. 5.2.2-5 angegebene Bezeichnung.

5.2.2.3. Konformationsisomerie. Lassen sich unterschiedliche Moleküle allein durch Rotation um Einfachbindungen ineinander überführen, so sind sie konformationsisomer zueinander. Von den unendlich vielen Konformationen eines Moleküls sind im allgemeinen nur

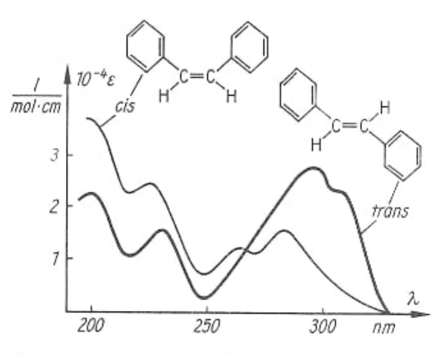


Abb. 5.2.2-4 Absorptionsspektren von cis- und trans-Stilben

Abb. 5.2.2-5 *E-Z*-Isomerie

wenige energetisch begünstigt. Die an sich freie Drehbarkeit um Einfachbindungen unterliegt demnach bestimmten Einschränkungen. Bei der Drehung sind *Rotationsbarrieren* zu überwinden. Abb. 5.2.2-6 zeigt die Abhängigkeit der

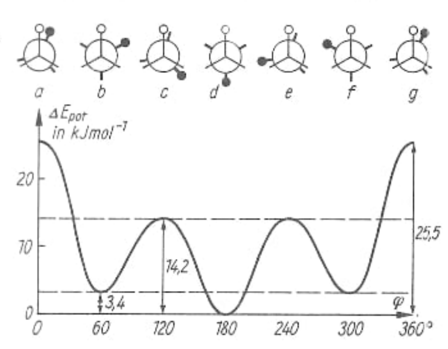


Abb. 5.2.2-6 Energieunterschiede der Konformationsisomeren des *n*-Butan in Abhängigkeit vom Drehwinkel φ um die C—C-Einfachbindung; *Punkte* in *a* bis *g* symbolisieren CH_3-Gruppen

potentiellen Energie des *n*-Butans vom Drehwinkel. Die ausgezeichneten Konformationen sind in einer *Newmanprojektion* veranschaulicht. Das Molekül wird dabei in Richtung der Bindung zwischen den mittleren C-Atomen betrachtet, alle anderen Bindungen sind in die Ebene projiziert. Der Schnittpunkt zwischen den drei durchgezogenen Bindungen symbolisiert das vordere, der Kreis das hintere C-Atom. Durch ein Minimum der potentiellen Energie sind die *gestaffelten Konformationen b, d* und *f* gekennzeichnet. Konformation *d* mit größter räumlicher Entfernung der Methylgruppen ist am stabilsten. Den *verdeckten* oder *ekliptischen Konformationen a, c* und *e* entsprechen Maxima der potentiellen Energie; *n*-Butan liegt bei Raumtemperatur etwa zu 80% in der Konformation *d* und zu 20% in den Konformationen *b* bzw. *f* vor.

Bei den ringförmigen gesättigten Kohlenwasserstoffen, den *Cycloalkanen* C_nH_{2n}, haben die C-Atome infolge der sp^3-Hybridisierung ebenfalls das Bestreben, die Tetraederstruktur zu wahren. Mit Ausnahme des notwendigerweise ebenen Cyclopropans sind die Cycloalkane nicht planar gebaut. Im Unterschied zu anderen ringförmigen Kohlenwasserstoffen gibt es bei Cyclohexan C_6H_{12} nahezu spannungsfreie *Konformere* (Abb. 5.2.2-7). Die stabilste Form ist die *Sesselform*. *Wannen-* und *Twistform* weisen infolge ekliptischer Partialkonformationen sowie Abstoßungskräften zwischen 1,4-ständigen H-Atomen eine 23,0 bzw. 19,9 $kJmol^{-1}$ höhere Energie auf. Im Cyclohexan liegen bei

Raumtemperatur mehr als 99,9% in der Sesselform vor. Twist- und Wannenform können z. B. in bi- oder polycyclischen Systemen fixiert sein.

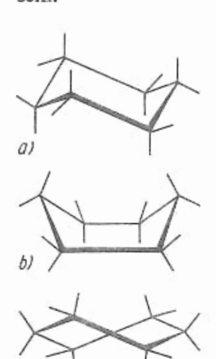


Abb. 5.2.2-7 Konformere von Cyclohexan; a) Sesselform, b) Wannenform, c) Twistform

5.2.3. Moleküle mit ungewöhnlichen Struktur- und Bindungsverhältnissen

5.2.3.1. Edelgasverbindungen. Platinhexafluorid PtF_6 reagiert bei Raumtemperatur mit Sauerstoff unter Bildung des Salzes $[O_2^+][PtF_6^-]$. Aufgrund des hohen 1. Ionisierungspotentials des molekularen Sauerstoffs, welches mit 1170 $kJmol^{-1}$ identisch mit dem des Xenons ist, schlußfolgerte Bartlett, daß eine analoge Reaktion zwischen Xenon und PtF_6 eintreten sollte. Dies führte im Jahre 1962 zur Synthese der ersten Edelgasverbindung, dem $Xe[PtF_6]$. Inzwischen ist die Chemie des Xenons relativ gut untersucht. Es wurden zahlreiche Verbindungen synthetisiert, z. B. XeF_4, XeF_6, $XeOF_4$, XeO_3. Von Helium, Neon und Argon konnten bisher keine stabilen Verbindungen hergestellt werden, was bei den hohen Ionisierungspotentialen dieser Edelgase auch sehr unwahrscheinlich wäre. Ein nützliches Hilfsmittel für das Verständnis der Elektronenstrukur von Edelgasverbindungen ist das Konzept der *Dreizentrenbindung*. Es soll am Beispiel des linearen XeF_2-Moleküls veranschaulicht werden (Abb. 5.2.3-1). Aus der Wechselwirkung eines besetzten 5*p*-Orbitals des Xenons mit zwei halbbesetzten 2*p*-Orbitalen der F-Atome resultieren drei Molekülorbitale. Von den vier in das MO-Schema einzubauenden Elektronen besetzen nur zwei ein bindendes MO. Die anderen beiden Elektronen sind nichtbindend, die Bindungsenergie der Xe—F-Bindung ist demnach geringer als die einer normalen Zweizentren-Zweielektronen-Bindung.

5.2.3.2. Borhydride. Bor bildet mit Wasserstoff eine Vielzahl von Verbindungen mit ungewöhnlicher Stöchiometrie. Schon das einfachste Borhydrid ist nicht BH_3, sondern *Diboran*, B_2H_6 (Abb. 5.2.3-2). Wie die anderen Borhydride ist Diboran eine *Elektronenmangelverbindung*. Die acht B—H-Bindungen im B_2H_6 werden durch nur 12 Valenzelektronen, d. h. sechs Elektronenpaare, vermittelt. Auch hier ist das Konzept der Dreizentrenbindung anwendbar: Die beiden etwa sp^3-hybridisierten B-Atome bilden normale Zweizentren-Zweielektronen-Bindungen mit den endständigen H-Atomen. Die beiden BH_2-Bruchstücke sind durch zwei B—H—B-Dreizentrenbindungen untereinander verbunden. Da für jede Dreizentrenbindung nur zwei Elektronen zur Verfügung stehen, ist im MO-Schema (↑ Abb. 5.2.3-1) nur das bindende MO besetzt. Es resultiert eine stabile Elektronenkonfiguration.

5.2.3.3. Übergangsmetall-π-Komplexe mit ungesättigten organischen Molekülen. Um 1830 beschrieb der dänische Apotheker Zeise eine Verbindung der Zusammensetzung $PtCl_2C_2H_4$ sowie Salze des Ions $[C_2H_4PtCl_3]^-$. Diese Substanzen, deren Struktur erst in neuerer Zeit völlig gesichert wurde, sind die ersten einer Vielzahl metallorganischer Verbindungen von Übergangsmetallen. Abb. 5.2.3-3 veranschaulicht schematisch die Bindung in Olefinkom-

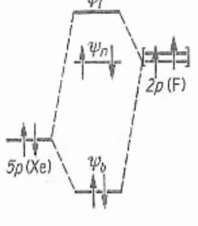


Abb. 5.2.3-1 Dreizentrenvierelektronenbindung in XeF_2

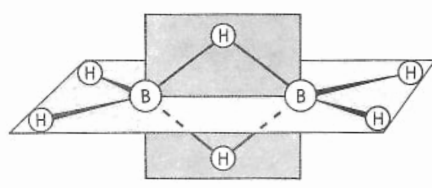


Abb. 5.2.3-2 Struktur des Diborans B_2H_6, Punktgruppe D_{2h}

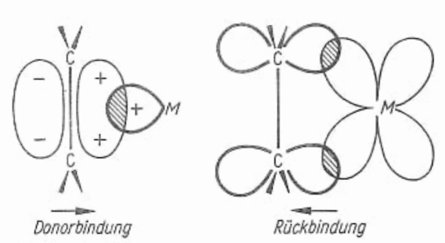


Abb. 5.2.3-3 Bindungsvorstellungen für Olefinkomplexe

plexen: Durch Überlappung eines besetzten π-Molekülorbitals des Olefins mit einer Akzeptorfunktion des Metallatoms vom σ-Typ wird Elektronendichte vom Ligand zum Metall übertragen, durch sogenannte Rückbindung zwischen besetzten Metall-*d*-Funktionen und antibindenden Bahnfunktionen des Olefins wird Elektronendichte in umgekehrter Richtung verschoben. Im Einklang mit den Bindungsvorstellungen wird eine Zunahme des C—C-Abstands und Abnahme der Frequenz der C—C-Valenzschwingung des Olefins bei der Komplexbildung beobachtet.

Auch aromatische Verbindungen wie Benzen oder das Cyclopentadienyl-Anion $C_5H_5^-$ sind zur Bildung von π-Komplexen mit Übergangsmetallen befähigt. Wegen ihrer Struktur werden diese Komplexe als Sandwichverbindungen bezeichnet. Ein intensiv untersuchter Vertreter ist das 1951 erstmals synthetisierte *Ferrocen* $(\pi\text{-}C_5H_5)_2Fe$. Neuerdings gelang sogar die Synthese von *Tripeldecker-Sandwich-Verbindungen* mit zwei gleichen oder auch unterschiedlichen Übergangsmetallatomen. Als erster Tripeldecker wurde 1972 das Tris-(η-Cyclopentadienyl)-dinickelkation beschrieben (Abb. 5.2.3-4).

5.2.3.4. Metallcarbonyle und Metall-Metall-Bindungen. Metallcarbonyle sind Verbindungen der *d*-Elemente mit Kohlenmonoxid. Es besteht eine systematische Beziehung zwischen der Zusammensetzung der Carbonyle und der Elektronenzahl, die dem Metallatom zur Erlangung der nächstfolgenden Edelgaskonfiguration fehlt. CO ist in der Lage, ein freies Elektronenpaar zur Bindung an ein Element mit Elektronenlücke zu verwenden. So erreichen Cr, Fe und Ni durch Koordination von sechs, fünf bzw. vier CO-Molekülen die Elektronenkonfiguration des Edelgases Krypton.

Die M—CO-Bindung ist überraschend stabil, sie hat ein sehr geringes Dipolmoment, die Bindungsordnung liegt zwischen eins und zwei. Die Bindungsverhältnisse ähneln denen in Olefinkomplexen (↑ Abb. 5.2.3-3). Durch Überlappung eines leeren σ-Orbitals des M-Atoms mit einem besetzten σ-Orbital des CO entsteht eine M—C—σ-Bindung. Dieser ist eine π-Bindung überlagert, die durch Überlappung besetzter *d*-Orbitale des Metalls mit dem leeren π^*-Molekülorbital von CO entsteht.

Bei Übergangsmetallen mit ungerader Ordnungszahl wird die Edelgaskonfiguration der Metallatome durch Bildung von mehrkernigen Komplexen mit CO-Brücken und M—M-Bindungen erreicht. In der in Abb. 5.2.3-5 gezeigten Verbindungsreihe wird die Verringerung der Zahl der CO-Liganden durch eine wachsende Zahl von Co—Co-Bindungen ausgeglichen. Es kommt zur Ausbildung von *Metallclustern*. Über die Bestimmung der M—M-Bindungslängen ist eine Ermittlung der Bindungsordnung möglich. Im $Co_4(CO)_{12}$ liegt z. B. eine Co—Co-Einfachbindung vor. Manchmal treten auch nicht ganzzahlige Bindungsordnungen auf. Für die beiden isostrukturellen Ionen $Mo_6Cl_8^{4+}$ und $Nb_6I_8^{3+}$ ist die Mo—Mo-Bindungsordnung 1,00, die Nb—Nb-Bindungsordnung hingegen 0,79. Das Ergebnis der Berechnung stimmt mit experimentellen Daten (Abb.5.2.3-6) überein.

Gegebenenfalls werden auch Mehrfachbindun-

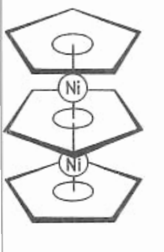


Abb. 5.2.3-4 Tripeldecker-Sandwich-Komplex

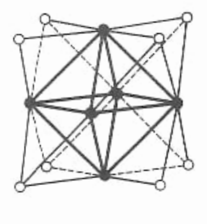
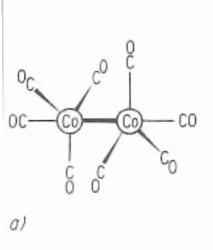
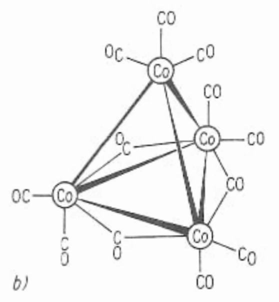


Abb. 5.2.3-6 Struktur von $Nb_6I_8^{3+}$- und $Mo_6Cl_8^{4+}$-Ionen





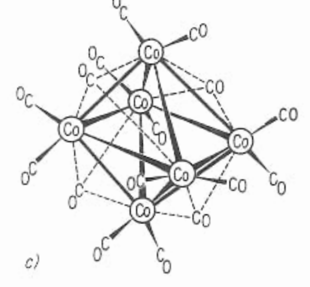


Abb. 5.2.3-5 Co-Carbonyle; a) $Co_2(CO)_8$, b) $Co_4(CO)_{12}$, c) $Co_6(CO)_{16}$

gen zwischen Metallen ausgebildet. Im $Re_2Cl_8^{2-}$-Ion liegt z. B. eine Re—Re-Vierfachbindung vor. Die Bindungsenergie dieser Bindung wird auf 1300 bis 1700 $kJmol^{-1}$ geschätzt und liegt damit höher als bei jeder anderen bekannten Bindung.

5.2.3.5. Ungewöhnliche organische Ringsysteme. Ungewöhnliche Ringsysteme haben Bedeutung für die Weiterentwicklung der Theorie der chemischen Bindung; die meist anspruchsvollen Synthesen bereichern auch die präparative organische Chemie (Abb. 5.2.3-7).

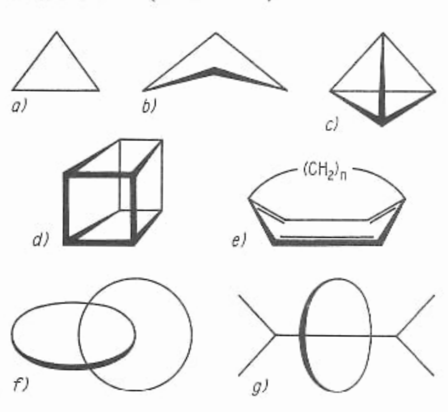


Abb. 5.2.3-7 Ungewöhnliche Ringsysteme; a) Cyclopropan, b) Cyclobutan, c) Tetrahedran (unbekannt), d) Cuban, e) [6]-Paracyclophan, f) Catenan, g) Rotoxan

Abb. 5.2.3-8 Kaliumkomplex eines Kronenethers

Der einfachste gesättigte zyklische Kohlenwasserstoff ist das *Cyclopropan* C_3H_6. Infolge der kleinen C—C—C-Winkel ist die Überlappung der Bindungsorbitale unvollkommen, es werden gebogene, sogenannte *Bananenbindungen* ausgebildet.

Cyclobutan liegt in Twistformen der Punktgruppe D_{2d} vor. Es ist stabiler als Cyclopropan. Der einfachste Kohlenwasserstoff mit der Struktur eines geometrischen Körpers, das *Tetrahedran* C_4H_4, konnte bisher nicht synthetisiert werden. Sein Valenzisomeres, Cyclobutadien, weist eine höhere Stabilität auf. Erfolgreich waren dagegen die Versuche zur Synthese von *Cuban* C_8H_8. Es besteht nur aus Cyclobutanringen, infolge der Raumstruktur ist die Bindungswinkelspannung höher als im Cyclopropan. Deformierte Bindungswinkel liegen auch bei den Cyclophanen vor. Der aromatische Sechsring im [6]-Paracyclophan ist z. B. nicht eben, die Bindungen zu den paraständigen C-Atomen sind um 22° aus der Ebene herausgedreht. *Catenane* und *Rotoxane* sind Ringsysteme höherer Ordnung, deren Zusammenhalt nicht ausschließlich durch chemische Bindungen hervorgerufen wird. Bei den Catenanen [*catena* lat., Kette] sind zwei Ringe kettengliedartig verknüpft, bei den Rotoxanen umschließt ein Kohlenwasserstoffring eine Kohlenwasserstoffkette mit sperrigen Endgruppen.

Die Anfang der siebziger Jahre erstmals synthetisierten *Kronenether* sind zur komplexen Bindung von Alkalimetallen befähigt. Sie enthalten heterocyclische Ringe mit O-Atomen. $KMnO_4$ löst sich z. B. in Gegenwart eines 18-gliedrigen Kronenethers in Benzen unter Bildung des in Abb. 5.2.3-8 dargestellten Salzes mit komplexem Kation.

5.3. Angeregte Moleküle und Molekülspektren

5.3.1. Energiezustände von Molekülen

Wenn man von Wechselwirkungseffekten geringer Bedeutung zunächst absieht, setzt sich die gesamte Energie eines Moleküls additiv aus der Translations- und der Rotationsenergie der Bewegung des ganzen Moleküls, aus der Energie von Schwingungen der Atomkerne gegeneinander und aus der Energie aller zum Molekül gehörenden Elektronen zusammen. Die Translationsenergie kann sich stetig ändern.

Für Rotationen, Schwingungen und Elektronenanregung existieren nur diskrete erlaubte Energiezustände.

5.3.1.1. Rotation der Moleküle. Nach den Gesetzen der klassischen Mechanik erfolgt die freie Rotation eines festen Körpers um eine Achse durch den Schwerpunkt. Bezüglich der dabei auftretenden Trägheitsmomente Θ zeichnen sich drei aufeinander senkrecht stehende Achsen, die *Hauptträgheitsachsen*, aus: eine Achse X mit Θ_X, dem kleinsten, eine Achse Z mit Θ_Z dem größten Wert von Θ und eine Achse Y mit einem mittleren Wert Θ_Y, festgelegt durch ihre Richtung. Wirft man eine in rasche Drehung versetzte Schachtel in die Luft, so lassen sich die Drehachsen X und Z mit minimalem Θ_X und maximalem Θ_Z finden, da deren Richtungen im Raum fest bleiben, während andere Achsen schwanken. Jede Drehbewegung läßt sich rechnerisch in Drehungen um die drei Hauptträgheitsachsen zerlegen.

Die *Rotationsenergie* $E_{rot} = \Theta\omega^2/2$ wird durch das Trägheitsmoment Θ und durch die Winkelgeschwindigkeit ω bestimmt. Bei einem Molekül aus 2 Atomen A und B sind die Massen m_A und m_B in den Atomkernen konzentriert. Die Bei-

träge der Elektronenhülle sind zu vernachlässigen. Da sich m_A, m_B im Kernabstand r_{AB} voneinander befinden, ist $\Theta = m_A m_B r_{AB}^2/(m_A + m_B)$. Bei diesen Molekülen erfolgt die Rotation um eine Achse, die senkrecht auf der Verbindungslinie der Kerne steht und durch den Massenmittelpunkt geht (Abb. 5.3.1-1). Die quantenmechanische Berechnung ergibt

$$(1) \quad E_{rot} = \frac{\hbar^2}{2\Theta} \cdot J(J+1) = hcBJ(J+1)$$

für die diskreten Niveaus der Rotationsenergie. Die Rotationsquantenzahl J kann die Werte $J = 0, 1, 2, \ldots$ annehmen; der Abstand der Niveaus wächst danach mit wachsender Quantenzahl J. Die Kreisfrequenz $\omega = \sqrt{2E_{rot}/\Theta} = (\hbar/\Theta)\sqrt{J(J+1)}$ kann als *Rotationsgeschwindigkeit* in Umdrehungen je Sekunde aufgefaßt werden, $B = \hbar/(4\pi c\Theta)$ ist die *Rotationskonstante* des betreffenden Moleküls. Für das HCl-Molekül mit $\Theta = 2{,}11 \cdot 10^{42}$ gm² hat der erste Rotationszustand, der auf den Grundzustand folgt, eine Drehzahl von $0{,}87 \cdot 10^{12}\,s^{-1}$, der zweite eine von $1{,}52 \cdot 10^{12}\,s^{-1}$, d. h. rund eine Billion Umdrehungen je Sekunde. Die Schwingungen der Atomkerne gegeneinander, die bei höheren Temperaturen einsetzen, haben eine etwa 100mal größere Frequenz.

Zu jedem Wert von J gibt es $(2J+1)$ verschiedene Realisierungsmöglichkeiten, d. h. wellenmechanisch: zu jeder Energie E_{rot}, die einer Drehimpulsquantenzahl zugeordnet ist, gibt es $(2J+1)$ verschiedene *Eigenfunktionen*, und die von ihnen beschriebenen Zustände unterscheiden sich nur durch die Lage der Rotationsachsen im Raum. Diese Lage wird durch eine weitere Quantenzahl m angegeben, die alle Werte von $-J$ über 0 bis $+J$ annehmen kann, in der Energieformel aber nicht auftritt. Man spricht von einer $(2J+1)$fachen Entartung und meint damit, daß $(2J+1)$ verschiedene Zustände zur gleichen Energie gehören. Die Trägheitsmomente der meisten zweiatomigen Moleküle liegen zwischen dem 10- und 100fachen von 10^{-42} gm². Bei Zimmertemperatur entfallen etwa 2400 Jmol⁻¹ auf die Molekülrotation, das entspricht Rotationsquantenzahlen zwischen ungefähr 10 und 30.

In die Energieberechnung bei mehratomigen Molekülen gehen die drei Hauptträgheitsmomente ein; für das H_2O-Molekül haben sie die Werte der Abb. 5.3.1-2. Da bei der Rotation der Moleküle keine einfache Überlagerung der Teilbewegungen erfolgt, kann die gesamte Rotationsenergie nicht als Summe von drei einzelnen Beiträgen dargestellt werden; es tritt eine komplizierte Kopplung der Bewegungen ein, wie es die Mechanik für den Kreisel lehrt. Für den allgemeinen Fall von drei verschiedenen Hauptträgheitsmomenten läßt sich überhaupt keine einheitliche Energieformel angeben, sondern nur für Spezialfälle mit teilweise gleichen Hauptträgheitsmomenten. Für die Beschreibung der Rotation kann man die Moleküle näherungsweise als *starre Körper* betrachten; man bezeichnet sie daher oft als Kreisel und unterscheidet lineare Moleküle, Kugelkreisel, symmetrische und asymmetrische Kreisel.

In *linearen Molekülen*, z. B. CO_2, CS_2, HC≡CH, HC≡C—C≡CH, N≡C—C≡C—C≡N, $HgCl_2$, $CdCl_2$, $ZnCl_2$, H—C≡N, Cl—C≡N, HgClBr, gibt es wie bei den zweiatomigen Mo-

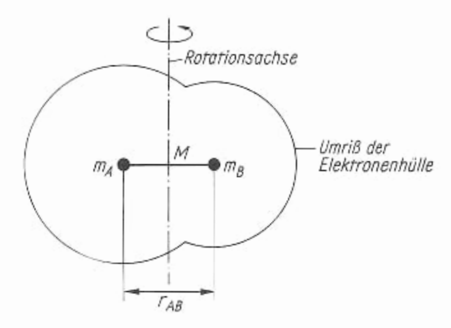


Abb. 5.3.1-1 Rotation eines zweiatomigen Moleküls (Größenverhältnisse für CO); r_{AB} Abstand der beiden Atomkerne A und B; M Massenmittelpunkt

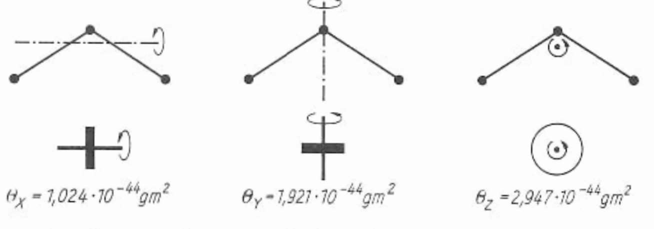


Abb. 5.3.1-2 Hauptachsen und Hauptträgheitsmomente für die Rotation des H_2O-Moleküls mit äquivalenten mechanischen Schwungradkreiseln zur besseren Vorstellung

lekülen nur ein wirksames Trägheitsmoment, denn das Trägheitsmoment um die Längsachse des Moleküls hat den Wert Null. Die Energieberechnung erfolgt wieder nach 5.3.(1), und das Trägheitsmoment $\Theta = \sum m_x r_x^2$ ist die Summe aller Atommassen m_x, multipliziert mit den Quadraten ihrer Abstände r_x vom Schwerpunkt.

Für *Kugelkreisel*, z. B. CH_4, CCl_4, P_4, SF_6, UF_6, $Ni(CO)_4$, $(CH_3)_4Pb$, ergibt sich wie bei einer Kugel für jede Achse durch den Mittelpunkt das gleiche Trägheitsmoment, es gilt wieder Formel 5.3.(1).

Symmetrische Kreisel, z. B. BF_2, C_6H_6, $CHCl_3$, C_2H_6, CH_3Cl, NH_3, haben eine ausgezeichnete *Figurenachse*, um die mindestens drei Atome oder Atomgruppen regelmäßig angeordnet sind und mit der eine Hauptträgheitsachse zusammenfällt. Das Trägheitsmoment ist maximal oder minimal, je nachdem, ob sich die Verteilung der schwereren Atome mehr einer Ebene oder einer Geraden nähert. Für die beiden übrigen Hauptachsen ergibt sich in jeder Lage das gleiche Trägheitsmoment. Es sind Formeln erforderlich, die zwei Trägheitsmomente und zwei verschiedene Quantenzahlen enthalten.

5.3.1.2. Schwingungen der Moleküle. Die Atomkerne schwingen im Molekül gegeneinander, wie eine Masse an einer Feder um einen festen Punkt schwingt. Ist die Federkraftkonstante f ein Maß für die elastischen Eigenschaften der Feder, so ist bei einer harmonischen Schwingung die rücktreibende Kraft $-F = f\Delta l$ der Dehnung oder Stauchung Δl der Feder proportional, und die Frequenz ist $\nu = \frac{1}{2\pi}\sqrt{\frac{f}{m}}$, im Fall eines leichten Atoms der Masse m, das gegen einen schweren Molekülrest schwingt, der näherungsweise in Ruhe bleibt. In einem zweiatomigen Molekül (Abb. 5.3.1-3) dagegen schwingen die Atome mit den Massen m_1 und m_2 gegen den Massenmittelpunkt M, und für m ist die *reduzierte Masse* $m = m_1 m_2/(m_1 + m_2)$ einzusetzen. Die Kraftkonstanten für Moleküle

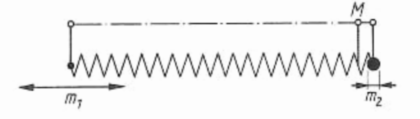


Abb. 5.3.1-3 Schwingung zweier verschiedener Massen m_1, m_2 gegen den Massenmittelpunkt M als Modell der Schwingung des zweiatomigen Moleküls, z. B. von HF

können aus den beobachteten Frequenzen berechnet werden und stehen in guter Übereinstimmung mit den Erfahrungen der Chemie über die Festigkeit der chemischen Bindungen; z. B. sind die Zahlenwerte für die Moleküle H—Cl, O=O und N≡N, d. h. für *Einfach-*, *Doppel-* und *Dreifachbindung*: $f = 470\ \mathrm{Nm^{-1}}$, $f = 1130\ \mathrm{Nm^{-1}}$ und $f = 2220\ \mathrm{Nm^{-1}}$. Allerdings ist der obige Ansatz für $-F$, der für eine harmonische Schwingung gilt, für Moleküle nur eine mangelhafte Näherung. Für die Verschiebung eines Atoms aus seiner Ruhelage r_0 zum Nachbaratom A hin ist nach $-F = f\Delta l$ die Energie $E = f(\Delta l)^2/2$ aufzuwenden; diese Funktion hat den Verlauf einer Parabel und auch für den Atomabstand $r = 0$ einen endlichen Wert, während in Wirklichkeit die Energie unendlich groß wird, weil sich die Atome nur in beschränktem Maß durchdringen können. Bei einer Bewegung von A weg, d. h. beim Auseinanderschwingen, muß dagegen bei stärkerer Dehnung die chemische Bindung geschwächt und schließlich zerrissen werden, so daß die *Dissoziationsenergie D* die obere Genze ist, in die die wahre Energiekurve einmünden muß. In der Nähe des Gleichgewichtsabstands r_e ist die harmonische Näherung brauchbar.

Die wellenmechanische Behandlung harmonischer Schwingungen zweiatomiger Moleküle liefert ebenso wie bei den Rotationen diskrete Energieniveaus $E_{vib} = h\nu(1/2 + v)$. In dieser Gleichung ist ν die Schwingungsfrequenz und v die Schwingungsquantenzahl, die ebenso wie J nur ganzzahlige Werte 0, 1, 2, ... annehmen kann. Im Unterschied zur Rotation ist hier der Abstand aufeinanderfolgender Energieniveaus gleich, die Niveaus sind äquidistant. Die Zahlenwerte für ν sind rund 100mal größer als die für Rotationsfrequenzen. Die Energiestufen sind so groß, daß die Moleküle bei Zimmertemperatur fast durchweg im Grundzustand $v = 0$ vorliegen.

Der Dissoziationssenergie entspricht gewöhnlich eine Schwingungsquantenzahl $v = 10$ bis $v = 40$ oder weit mehr für langsam schwingende schwere Atome. Im Grundzustand hat das Molekül noch Schwingungsenergie $1/2(h\nu)$, die ihm durch keine Abkühlung entzogen werden kann; die Atome der Moleküle führen noch am absoluten Nullpunkt Nullpunktschwingungen aus. Ein schwingungsloser Grundzustand würde bedeuten, daß die Atome in ihrem Gleichgewichtsabstand in Ruhe nebeneinander liegen, daß man demnach den Ort und den Impuls für beide Atomkerne zugleich genau angeben könnte; das aber würde der Heisenbergschen Unschärfebeziehung widersprechen. Die Strekken, die bei den einzelnen Energiestufen zwischen den von der Energiekurve gesetzten Grenzen liegen (Abb. 5.3.1-4), bedeuten die Amplituden vom Atom B, wenn man A als räumlich fixiert annimmt. Bei der wellenmechanischen Ableitung ergeben sich statt Amplituden Hauptmaxima der Aufenthaltswahrscheinlichkeit, die den Umkehrpunkten bei der Schwingung

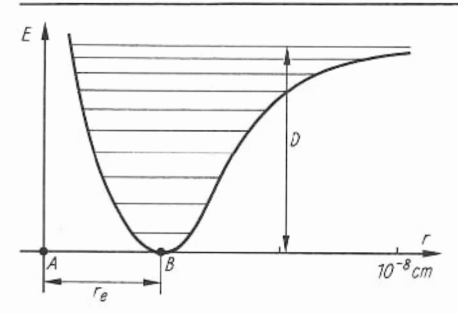


Abb. 5.3.1-4 Energiebedarf für die Dehnung oder Kompression einer chemischen Bindung; r_e Gleichgewichtsabstand der Atome *A* und *B*, *D* Dissoziationsenergie; die *horizontalen Stufen* stellen die quantenmäßigen Niveaus der Schwingungsenergie dar

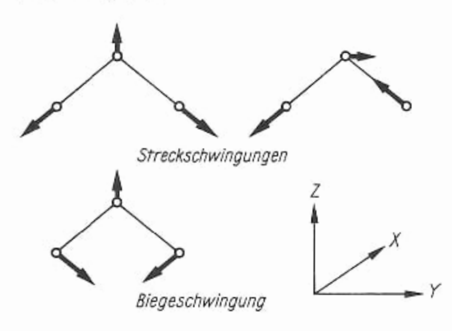


Abb. 5.3.1-5 Die drei Grundschwingungen eines dreiatomigen, symmetrischen, gewinkelten Moleküls; die Amplituden sind etwa 100mal zu groß gezeichnet

entsprechen, bei etwas kleineren Atomabständen.

Ein Molekül aus n Atomen hat $3n - 6$ bzw. bei linearer Anordnung $3n - 5$ innere Schwingungen. Die Valenz- oder Streckschwingungen erfolgen in Richtung der chemischen Bindung, Deformations- oder Biegeschwingungen verändern die Winkel zwischen benachbarten Bindungen (Abb. 5.3.1-5). Beim Wasser z. B. ist eine Deformationsschwingung möglich, und zwei Freiheitsgrade verbleiben für Valenzschwingungen. Dabei kann aber nicht die eine nur von der Schwingung der linken OH-Valenz und die andere allein von der rechten herrühren, da dann der Massenmittelpunkt nicht in Ruhe bliebe. Er liegt beim H_2O-Molekül auf der *Z*-Achse in der Nähe des O-Atomkerns und bleibt unverändert, wenn die linke OH-Valenz mit der gleichen Phase, Frequenz und Amplitude schwingt wie die rechte und wenn weiter die Amplituden auf das O-Atom und die H-Atome so verteilt sind, daß die Massenverlagerung infolge der Bewegung des Sauerstoffs durch die der Wasserstoffatome gerade ausgeglichen wird. Die Lage des Massenmittelpunkts bleibt andererseits auch erhalten, wenn die beiden OH-Valenzen mit genau entgegengesetzter Phase schwingen. Damit bringen sie selbst die Massenverlagerung in der *Z*-Richtung zum Verschwinden. Nur die *Y*-Komponente muß durch eine entgegengesetzte, kleine Bewegung des O-Atoms kompensiert werden. Auch bei einer Biegeschwingung genügt eine kleine Bewegung des O-Atoms. Für Atomkernabstände von rund 10^{-10} m betragen die Amplituden $\approx 10^{-12}$ m. Eine Schwingungsperiode ist abgeschlossen, wenn die Atome vom Ort der Pfeilspitze (↑ Abb. 5.3.1-5) über ihre Ruhelage hinaus auf die entgegengesetzte Seite und wieder in die Ausgangslage zurückgelangt sind. Die Zeiten für diesen Schwingungsvorgang betragen für die schnellere der beiden Streckschwingungen $0{,}89 \cdot 10^{-14}$ s, für die zweite $0{,}91 \cdot 10^{-14}$ s und für die Biegeschwingung $2{,}09 \cdot 10^{-14}$ s. Die Werte für die entsprechenden Energien im Wellenzahlmaßstab, die Termwerte, sind 3756 cm^{-1}, 3654 cm^{-1} und 1596 cm^{-1}. So wird in der Spektroskopie die Energie durch die Wellenzahl $\tilde{\nu} = 1/\lambda$ als reziproker Wert der Wellenlänge λ einer energiegleichen Strahlung angegeben.

Das Beispiel des H_2O lehrt die Grundgesetze der Schwingungen beliebig vielatomiger Moleküle: Es bewegen sich alle Atome mit gleicher Frequenz und Phase (bis auf Sonderfälle, in denen ein Atom in Ruhe bleiben kann, wenn es z. B. selbst im Massenmittelpunkt liegt). Die Schwingungen sind gekoppelt, wie bei Pendeln oder elektrischen Schwingkreisen. Im Resonanzfall, bei Kopplung von Oszillatoren gleicher Eigenfrequenz (wie den 2 OH-Valenzen bei H_2O) gilt

$$(2) \quad \nu_{1,2} = \nu_0 \sqrt{\frac{1 \pm m_1 \cos \alpha}{m_1 + m_2}}$$

wenn ν_0 die Frequenz des ungekoppelten Oszillators, m_1 die Masse des Endatoms, m_2 die des Mittelatoms, über das die Kopplung erfolgt, und α den Valenzwinkel bedeuten. Die beiden Frequenzen ν_1 und ν_2 sind nach oben und nach unten gegen den Wert ν_0 für den freien zweiatomigen Oszillator verschoben. Diese Erscheinung heißt Resonanzaufspaltung und kehrt beim Zusammentreten gleicher Valenzen in einem Molekül immer wieder. Aus den Werten für H_2O erhält man für die einzelne OH-Valenz $\tilde{\nu}_0 = 3705$ cm^{-1}, was mit dem Wert der Wellenzahl des OH-Ions ≈ 3650 cm^{-1} oder dem Wert 3570 cm^{-1} des OH-Radikals, das in Flammen vorkommt, vergleichbar ist. Da die Massen in diesen Beispielen gleich sind, beruhen die Wellenzahlunterschiede auf etwas verschiedenen Kraftkonstanten; die OH-Bindung im Wassermolekül ist als fester zu betrachten.

Aus Versuchen mit Pendeln kennt man auch stark gekoppelte Schwingungen verschiedener Frequenz. Entsprechend tritt in Molekülen noch eine Beeinflussung zwischen Valenz- und Deformationsschwingungen ein. Beim Wassermolekül kann nur die symmetrische Valenzschwingung mit der Biegeschwingung in Wechselwirkung treten; für die antisymmetrische ist das nicht möglich, weil die Bahnen der einen und der anderen Schwingung für jedes Atom senkrecht aufeinander stehen und sich damit nicht beeinflussen. Nach der mathematischen Bezeichnung „normal" für „senkrecht zueinander" bezeichnet man die Grundschwingungen auch als *Normalschwingungen*.

Für 3654 cm^{-1} und die Biegeschwingung stimmen die Bewegungsrichtungen am O-Atom überein. Es wird Energie übertragen. Die isoliert gedachten Schwingungen mischen sich. Die Bewegungsrichtungen weichen um 3° von der Valenzrichtung bzw. von der Biegerichtung senkrecht dazu ab. In der Energie für $\tilde{\nu} = 3654\ cm^{-1}$ ist rund 1% Energie der Deformationsschwingung enthalten. Dagegen besteht die Energie der Schwingung für $\tilde{\nu} = 1596\ cm^{-1}$ aus 0,6% Valenz- und 99,4% Deformationsanteil. Bei enger benachbarten Wellenzahlen wird die Mischung stärker. Über diese festen Verhältnisse hinaus tritt keine weitere Beeinflussung ein. Streckschwingung und Biegeschwingung sind unter Einrechnung dieser Mischungsanteile betrachtet unabhängig, also in diesem Sinn Normalschwingungen.

Die Formen der Grundschwingungen wurden unter der Voraussetzung abgeleitet, daß der Massenmittelpunkt in Ruhe bleibt. Sie sind deshalb für symmetrische Moleküle (↑ 5.2.1.) übersichtlich. In *symmetrischen Normalschwingungen* stimmt die Schwingungsrichtung für symmetrische Punkte überein; in antisymmetrischen Normalschwingungen ist sie entgegengesetzt (Abb. 5.3.1-6). Die Schwingung $\tilde{\nu} = 3654\ cm^{-1}$ und ebenso die mit $\tilde{\nu} = 1596\ cm^{-1}$ beim H_2O-Molekül erfolgen symmetrisch zu jedem Symmetrieelement, d. h. totalsymmetrisch. $\tilde{\nu} =$ 3756 cm^{-1} dagegen ist antisymmetrisch gegen die senkrechte Spiegelungsebene und gegen die zweizählige Achse, weil die Valenz links verlängert, aber rechts um den gleichen Betrag verkürzt wird. Dieser Vorgang bleibt aber symmetrisch zur Zeichnungsebene.

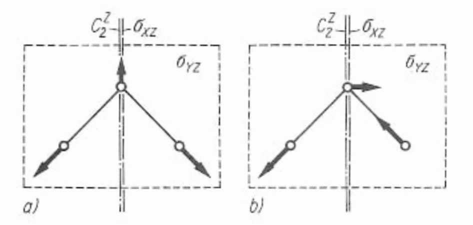


Abb. 5.3.1-6 Molekülschwingungen und Molekülsymmetrie; a) symmetrische Valenzschwingung, totalsymmetrisch; b) antisymmetrische Valenzschwingung, antisymmetrisch gegenüber C_2 und σ_{xz}, symmetrisch gegenüber σ_{yz}

Nach dem Verhalten gegenüber den Symmetrieelementen teilt man die Normalschwingungen in *Schwingungsrassen* ein. Schwingungen aus verschiedenen Rassen koppeln nicht untereinander, wie nahe benachbart die Frequenzen auch sein mögen. Das ist der allgemeinere Grund, weshalb z. B. bei H_2S die symmetrische und die antisymmetrische Valenzschwingung mit fast gleicher Wellenzahl erfolgen kann, ohne daß eine Resonanzaufspaltung eintritt, und weshalb die antisymmetrische Valenzschwingung bei H_2O keinen Deformationsanteil enthält. Für jede Punktgruppe lassen sich mittels der Gruppentheorie die Anzahl und Art der Rassen und damit die Form der in einem Molekül möglichen Schwingungen ableiten.

5.3.1.3. Elektronenanregung von Molekülen. Wie die Elektronenhülle der Atome kann auch das Elektronensystem der Moleküle durch Energiezufuhr in angeregte Zustände übergehen. Die Berechnungen der Energien auf der Grundlage der Wellenmechanik können – abgesehen vom praktisch unwichtigen H_2^+-Ion – nur durch Näherungsverfahren erfolgen (↑ 5.1.2.).

5.3.1.3.1. *Änderungen der Hauptquantenzahl.* Angeregte Zustände, in denen in Analogie zu den Atomen ein Valenzelektron auf eine weiter außen liegende Bahn gehoben wird, sind für die Moleküle von untergeordneter Bedeutung. Durch die Verbindungsbildung werden die Valenzelektronen fester gehalten als bei den Atomen. Führt man entsprechend größere Energien zu, so tritt gewöhnlich die Dissoziation ein; z. B. erfordern die Anregungszustände des H_2-Moleküls, für die ein Elektron die Hauptquantenzahl $n = 2$ annimmt, Energien von $\approx 1096\ kJ\ mol^{-1}$, während die Dissoziationsenergie nur rund 400 $kJ\ mol^{-1}$ beträgt. Derartige angeregte Moleküle sind nur im Gaszustand und meist nur kurze Zeit existenzfähig. Man kennt aber größere Moleküle, in denen die Änderung einer Hauptquantenzahl eine ganze Reihe von Niveaus erzeugen kann. Diese *Rydbergniveaus der Moleküle* lassen sich wie bei den Atomen durch $E = R/(n + a)^2$ darstellen, wobei R die Rydbergkonstante, a eine Konstante der Spektralserie und n eine positive ganze Zahl bedeuten.

5.3.1.3.2. *Weitere Anregungszustände zweiatomiger Moleküle.* Die weitaus größere Bedeutung haben Änderungen der Nebenquantenzahl oder Bahnquantenzahl, die bei Molekülen mit Λ bezeichnet wird und die mit der Form des Molekülorbitals zusammenhängt. Die von den Atomen her bekannten Gesetze kann man auf

zweiatomige oder lineare Moleküle ausdehnen, da sich gewissen Elektronen ein bestimmter Bahndrehimpuls um die Molekülachse zuschreiben läßt. Im Grundzustand fast aller stabilen Moleküle ist infolge der Absättigung der Bahndrehimpuls Null. Für die Buchstabensymbole $S, P, D, F, G, \ldots$ der resultierenden Bahnquantenzahl $L = 0, 1, 2, 3, 4, \ldots$ des Atoms (↑ 4.2.4.2.) verwendet man beim Molekül die entsprechenden griechischen Buchstaben Σ, Π, Δ, Φ, Γ, ...; wie beim Atom wird die aus den Elektronenspins folgende Multiplizität als linker oberer Index diesem Symbol angefügt. Bei symmetrischen Molekülen sind noch Angaben über das Verhalten ihrer Wellenfunktion notwendig, die als Indizes an die Termsymbole gesetzt werden. Das Symbol $^1\Sigma_g^+$, das für den Grundzustand fast aller zwei- und gleichatomiger Moleküle gilt, z. B. für H_2, N_2, Cl_2, bedeutet: Die Elektronen sind so angeordnet, daß kein magnetisches Moment, weder von der Bahnbewegung ($\Lambda = 0$, deshalb Σ) noch vom Spin (Multiplizität 1) übrigbleibt. Außerdem wird durch ^+_g angegeben, daß die Wellenfunktion bei allen Symmetrieoperationen das Vorzeichen beibehält. Unter den ersten Anregungszuständen des Wasserstoffs finden sich solche mit $^1\Pi$, in denen eines der Elektronen auf einer Bahn kreist, die einen Bahndrehimpuls hat; durch $^3\Sigma$, $^3\Pi$ wird angegeben, daß die Spins der beiden Elektronen nicht mehr entgegengesetzt, sondern parallel gerichtet sind. Für das Sauerstoffmolekül O_2 im Grundzustand gilt $^3\Sigma_g^-$, d. h., 2 Elektronen sind ungepaart und verursachen den Paramagnetismus des Sauerstoffs. Die Beseitigung der Parallelstellung der Spins zu $^1\Sigma_g^+$ erfordert eine Anregungsenergie von rund 160 kJ mol^{-1}.

5.3.1.3.3. *Leicht anregbare Elektronen in komplizierten Verbindungen.* Bei allen größeren und nicht hochsymmetrischen Molekülen muß man auf die genaue Klassifizierung der Elektronenterme verzichten. Es ist aber wichtig, daß qualitativ bereits für viele Struktureinheiten angegeben werden kann, ob eine Elektronenanregung leicht und relativ unabhängig vom Rest des Moleküls eintreten kann. Durch leichte Anregbarkeit zeichen sich alle π-Elektronen aus, z. B. die in C=C-Doppelbindungen. Sie sind auch von den σ-Elektronen in den übrigen Bindungen unabhängig, ihre Anregung kann das Molekül kaum zerstören. Für die Chemie hat deshalb die optische Untersuchung von π-Elektronensystemen die Hauptbedeutung. Bei den anorganischen Verbindungen zeichnen sich weiter Atome mit d- oder f-Elektronen durch leichte Anregbarkeit aus. Die Übergangselemente bilden deshalb farbige Verbindungen. Durch Koordination in Kristallgittern oder Komplexen werden ihre Energieniveaus beträchtlich verschoben und damit die Farben verändert.

5.3.1.4. **Gesamtenergie eines Moleküls.** Die bisher behandelten Energieanteile haben im Wellenzahlmaß die Größenordnung 0,1 bis 10 cm^{-1} für Rotationsenergie, 100 bis 5000 cm^{-1} für Schwingungsenergie und 10^4 bis 10^5 cm^{-1} für Elektronenanregungsenergie. Nach $\tilde{\nu} = 1/\lambda$ entsprechen den Rotationsenergien Wellenlängen von 100 bis zu 1 mm, die im Mikrowellenbereich liegen, der Schwingungsenergie Wellenlängen zwischen 2 und 100 µm im nahen und fernen Ultrarot und den Elektronenenergien Wellenlängen zwischen 0,1 und 1 µm im Bereich ultravioletter, sichtbarer und naher ultraroter Strahlung. Die hohen Werte für Rotations- und Schwingungsenergie treten allerdings nur auf, wenn die leichten Wasserstoffatome an der betreffenden Bewegung beteiligt sind. Durch graphische Darstellung der drei Energieanteile erhält man das Termschema (↑ Abb. 5.3.2-2). Es ist für mehratomige Moleküle wegen der Überlagerung der 3 Rotations- und der $3n - 6$ Schwingungsmöglichkeiten außerordentlich kompliziert.

Die bisher betrachteten Energiezustände der Moleküle bilden sich je nach der herrschenden Temperatur aus, indem durch Stöße zwischen den Molekülen eine Gleichverteilung der Energie angestrebt wird. Nicht im thermischen Gleichgewicht und dabei oft in sehr hohen Anregungszuständen befinden sich Moleküle, wenn sie aus chemischen Reaktionen hervorgehen. Eine besondere Bedeutung hat die Anregung durch elektromagnetische Strahlung, also durch Licht im weitesten Sinne.

5.3.2. Optische Übergänge in Molekülen

Unter dem Einfluß elektromagnetischer Strahlung können Übergänge zwischen den Energieniveaus der Moleküle erfolgen, falls die Bohrsche Frequenzbedingung $\Delta E = \hbar\omega$ erfüllt ist, d. h., wenn die Frequenz $\nu = \omega/(2\pi)$ der mit dem Molekül in Wechselwirkung tretenden Strahlung einer der im Termschema vorkommenden Energiedifferenzen ΔE entspricht. Beim Übergang in ein höheres Energieniveau wird Strahlung absorbiert und das Molekül angeregt, im umgekehrten Fall wird Strahlung emittiert. Da sich bei Zimmertemperatur praktisch alle Moleküle in ihren tiefsten Elektronen- und Schwingungszuständen befinden und nur Rotationszustände merklich angeregt sind, erfordert eine Emission vorhergehende Energiezufuhr. Entweder durch Absorptions- oder Emissionsexperimente erhält man das *Spektrum* eines Moleküls.

Es treten dabei nicht alle möglichen Übergänge wirklich auf; nur relativ wenige sind aufgrund von *Auswahlregeln*, die aus den Gesetzen der Wellenmechanik folgen, erlaubt, der Rest wird als *verboten* bezeichnet. Die Auswahlregeln gelten zunächst für die elektrische Dipolstrah-

lung, die man sich durch Änderungen, speziell Schwingungen eines elektrischen Dipols wie Radiowellen entstanden denkt; wegen magnetischer und höherer elektrischer Momente werden auch verbotene Übergänge beobachtet, aber mit wesentlich kleinerer Intensität. Einige Auswahlregeln beruhen auf bestimmten Symmetrien. Störungen der Symmetrie lassen die verbotenen Übergänge mit erheblicher Intensität erscheinen.

5.3.2.1. Rotationsübergänge. Ändert sich nur der Rotationszustand, so gilt für die Rotationsquantenzahl J die Auswahlregel $\Delta J = \pm 1$, d. h., es sind nur Übergänge zwischen unmittelbar benachbarten Termen möglich. Die entsprechende Energiedifferenz $\Delta E = E_{rot}(J+1) - E_{rot}(J)$ ergibt sich für zweiatomige Moleküle zu $\Delta E = \hbar^2 J/\Theta$, und für die Wellenzahl der emittierten Strahlung erhält man

$$(3) \quad \tilde{\nu} = \Delta E/(hc) = hJ/(4\pi^2 c\Theta) = 2BJ$$

mit $B = h/(4\pi c\Theta)$ und $\tilde{\nu} = \nu/c$. Mit wachsendem J nimmt $\Delta\tilde{\nu}$ in gleichen Schritten zu; man erhält daher als Rotationsspektrum eine Folge äquidistanter Linien. B liegt bei schweren Molekülen in der Größenordnung von 0,1 cm^{-1} entsprechend 10 cm Wellenlänge, bei leichten von rund 10 cm^{-1} oder 0,1 cm Wellenlänge. Die reinen Rotationsspektren können nur im Gaszustand beobachtet werden. Schon bei größerer Dichte werden infolge zwischenmolekularer Wechselwirkungen die Linien verbreitert. In Flüssigkeiten und Festkörpern verschmelzen sie zu *Banden.*

Bei Präzisionsmessungen stellt sich heraus, daß die Linien nicht genau äquidistant sind, sondern mit wachsendem J einander näher rücken. Dies ist auf den Einfluß der Zentrifugalkraft zurückzuführen, die bei schneller Rotation die Bindung merklich dehnt. Der Kernabstand und damit das Trägheitsmoment werden größer und infolgedessen B kleiner.

Die Rotationsspektren mehratomiger Moleküle sind wegen der komplizierten Bewegungsverhältnisse verwickelter und lassen sich nur mühsam analysieren.

5.3.2.2. Schwingungsübergänge. Für harmonische Schwingungen gilt die Auswahlregel $\Delta v = \pm 1$, d. h., es sind nur Übergänge zwischen unmittelbar benachbarten Termen erlaubt (↑ 5.3.2.1.). Da sich gleichzeitig auch der Rotationszustand ändert, erhält man ein *Rotationsschwingungsspektrum*, für das nun weiter die Auswahlregel $\Delta J = \pm 1$ gilt. Für das zweiatomige Molekül HCl z. B. sind die Schwingungsniveaus mit $v = 0$ und $v = 1$ zu betrachten; wegen des größeren Trägheitsmoments im angeregten Schwingungszustand ($B_0 \approx 10{,}0$ cm^{-1}; $B_1 \approx 10{,}4$ cm^{-1}) liegen die oberen Rotationsniveaus geringfügig enger als die unteren (Abb. 5.3.2-1). Zwischen ihnen existieren erlaubte Übergänge; der zwischen

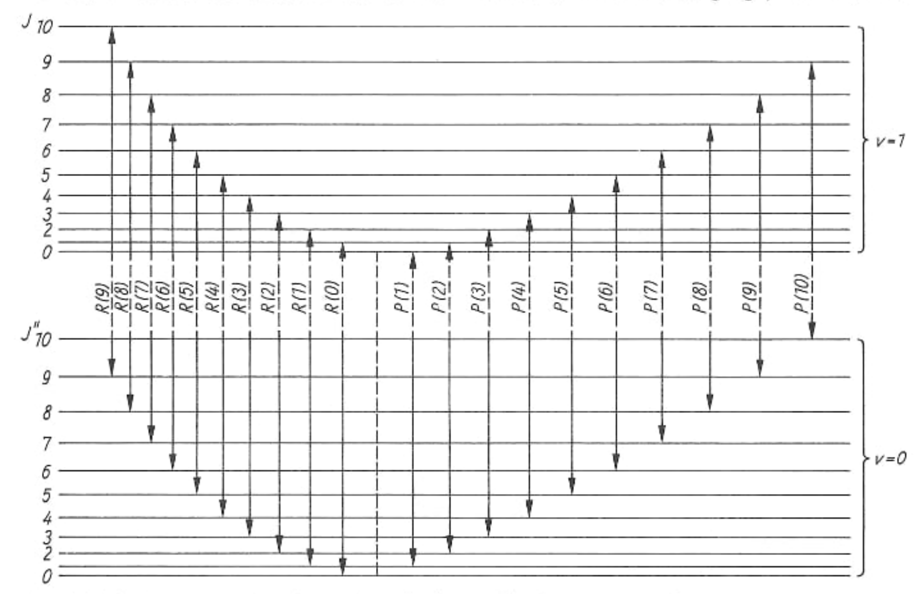


Abb. 5.3.2-1 Entstehen einer Rotationsschwingungsbande aus den Übergängen zwischen den Rotationstermen von HCl (Absorption); der kleinste J-Wert im R-Zweig ist 0 und im P-Zweig 1, der verbotene Übergang ist gestrichelt

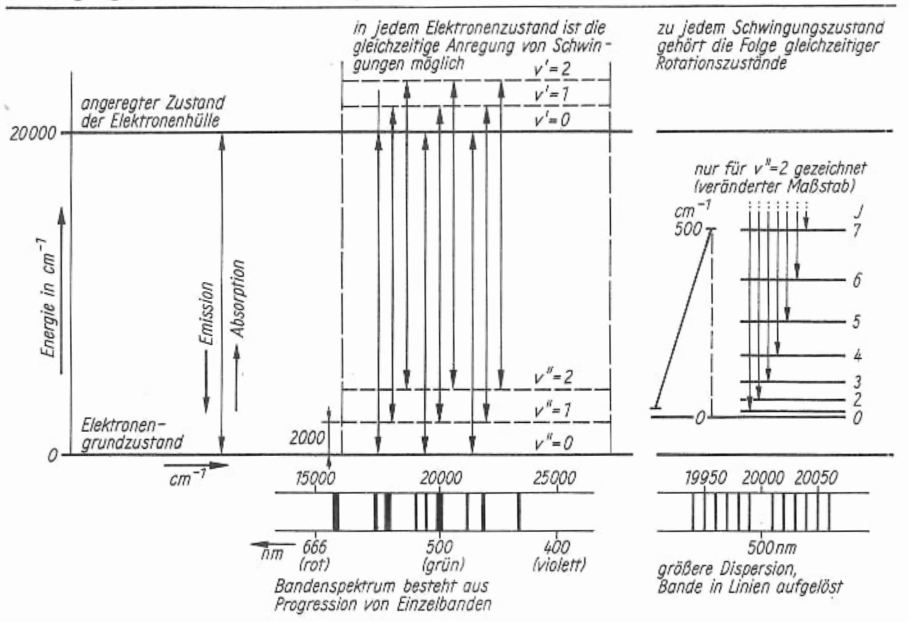


Abb. 5.3.2-2 Entstehung des Bandenspektrums durch Elektronenübergang und gleichzeitige Schwingungs- und Rotationsübergänge

$J' = 0$ oben und $J'' = 0$ unten ist verboten. Das zugehörige Absorptionsspektrum besteht aus zwei Zweigen eng benachbarter Linien, aus dem kurzwelligen *R*-Zweig, der den Übergängen $\Delta J = +1$, und aus dem langwelligen P-Zweig, der den Übergängen $\Delta J = -1$ entspricht.

Der Unterschied der Rotationskonstanten wirkt sich so aus, daß im *P*-Zweig die Linienabstände zunehmen und im *R*-Zweig abnehmen. Bildet man die Differenzen zweier Zustände nach 5.3.(3) unter Berücksichtigung von $\Delta J = 1$ und -1, verschiedener *B*-Werte und einer konstanten Schwingungswellenzahl $\tilde{\nu}_0$, so erhält man

$$(4)\quad \tilde{\nu}_R = \tilde{\nu}_0 + 2B' + (3B' - B'')\,J'' + (B' - B'')\,J''^2$$
$$\tilde{\nu}_P = \tilde{\nu}_0 - (B' + B'')\,J'' + (B' - B'')\,J''^2$$

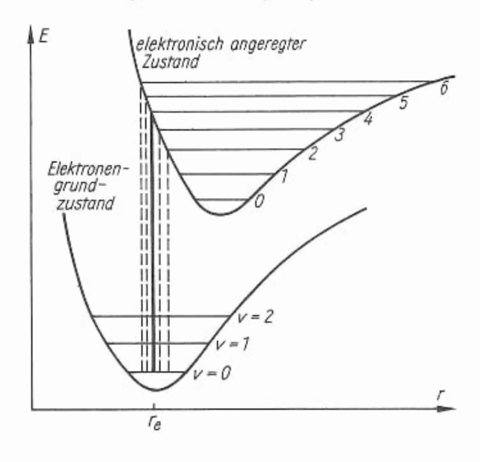
Die den oberen Zustand betreffenden Größen werden immer mit ′, die für den unteren mit ″ gekennzeichnet. Das quadratische Glied ist die Ursache der nichtkonstanten Linienabstände. Die bei den reinen Rotationsspektren erwähnte Zentrifugaldehnung ist hier wegen der Kleinheit des Effekts nicht wahrnehmbar.

Bei nichtlinearen Molekülen und bei solchen Schwingungen, bei denen ein lineares Molekül nichtlinear deformiert wird, sind auch Übergänge $\Delta J = 0$ erlaubt, die sich als besonderer *Q*-Zweig (wie man sich leicht bei Abb. 5.3.2-1 hinzudenkt) bei ν_0 und wenig kleineren Werten anhäufen. Der *Q*-Zweig ist mit normalen Spektrometern nicht in Linien auflösbar. Überhaupt läßt sich die Rotationsstruktur des Schwingungsübergangs bei mehratomigen Molekülen, die eine kleine Rotationskonstante haben oder bei denen sich verschiedene Rotationen überlagern, nicht auflösen. Man erhält dann aber infolge der Intensitätsverteilung noch eine charakteristische Bandenkontur mit zwei, drei oder vier Maxima. In Flüssigkeiten und bei festen Stoffen fehlt die Rotationsstruktur, und man erhält eine Bande, deren Maximum man als die Frequenz des Schwingungsübergangs ansieht.

5.3.2.3. Elektronenübergänge. Bei Elektronenübergängen ändert sich gleichzeitig auch der Schwingungs- und der Rotationszustand, so daß trotz bestimmter Auswahlregeln eine ungeheure Vielzahl von Linien auftritt und eine sehr breite Elektronenanregungsbande entsteht. Einigermaßen übersichtliche Verhältnisse ergeben sich bei den zweiatomigen Molekülen. Da es für sie nur eine Schwingungsfrequenz gibt, tritt eine ausgeprägte Schwingungsstruktur auf; die Elektronenanregungsbande zerfällt in ein *Bandensystem*, eine Reihe äquidistanter *Schwingungsbanden*. Im Prinzip sind dabei beliebige Änderungen der Schwingungsquantenzahl v möglich (Abb. 5.3.2-2).

Bei Untersuchung im Gaszustand und mit

genügender Auflösung zerfallen die Schwingungsbanden noch in einzelne Rotationslinien. Für die Quantenzahl J gilt dabei im allgemeinen die Auswahlregel $\Delta J = 0$ oder ± 1. Man erhält wieder R-, Q- und P-Zweig für jede Bande.



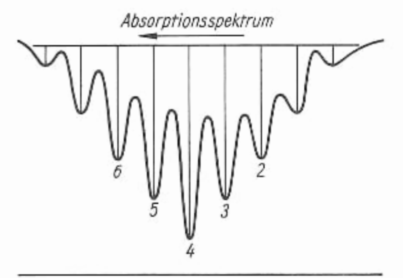


Abb. 5.3.2-3 Franck-Condon-Prinzip

Bei der Rotationsstruktur der Elektronenanregungsspektren ist in besonderem Maße zu beachten, daß die Trägheitsmomente im Anregungszustand nun wesentlich andere, meist größer sind. Das nach 5.3.(4) in J quadratische Glied kann bei großen Rotationsquantenzahlen das lineare Glied übertreffen, die Linienfolge kehrt im P-Zweig zu wieder größer werdenden Frequenzwerten um. An der Umkehrstelle entsteht eine Häufung von Einzellinien, ein *Bandenkopf*, im Unterschied zu dem dünneren anderen Ende der Bande, der *Abschattierung*.

Die Rotationslinien liegen eng benachbart. Ihr Abstand beträgt für das leichteste Molekül H_2 rund $20\,cm^{-1}$ und für zweiatomige Moleküle ohne Beteiligung von Wasserstoff einige cm^{-1}. Er ist für die Mehrzahl aller Moleküle viel kleiner. Man kann deshalb die einzelnen Rotationslinien meist nicht voneinander trennen. Bei den Absorptionsmessungen an Lösungen und flüssigen Stoffen ist überhaupt nicht mehr mit der Rotationsstruktur zu rechnen. Es verbleibt, weil die Moleküldrehungen nicht mehr frei erfolgen, nur noch eine strukturlose *Rotationsverbreiterung* der Schwingungsbanden.

Die Verteilung der Intensität auf die einzelnen Schwingungsbanden wird durch eine Art Auswahlregel, das *Franck-Condon-Prinzip*, bestimmt, nach dem die Elektronenübergänge zwischen denjenigen Schwingungszuständen am wahrscheinlichsten sind, für die die Atomkernabstände r (Abb. 5.3.2-3) übereinstimmen. Wegen der geringen Masse der Elektronen erfolgt die Änderung ihrer Anordnung so schnell, daß sich Lage und Geschwindigkeit der Kerne während eines Elektronensprungs nicht wesentlich ändern. Dabei bringen die Atomkerne die längste Zeit in der Nähe der Umkehrpunkte ihrer Schwingung zu, in denen ihre Geschwindigkeit Null ist. Je nachdem, ob das Minimum der Kurve der potentiellen Energie für den elektronisch angeregten Zustand bei kleinerem oder bei größerem Kernabstand r liegt als das Minimum der Kurve für den Grundzustand bei r_e, wird die Bindung des angeregten Elektrons verfestigt oder gelockert. Die Lockerung kann zu anschließender Dissoziation führen. Was zunächst Gleichgewichtsabstand r für den Grundzustand war, bleibt erhalten als der Kernabstand eines höheren Schwingungszustands. Für das Emissionsspektrum gilt umgekehrt, daß der Grundzustand am besten erreicht werden kann, wenn sich die Atome des Moleküls gerade mit einer Amplitude nähern, die dem angestrebten Gleichgewichtsabstand entspricht. Im Spektrum treten dann neben einem besonders begünstigten Übergang (starke senkrechte Linie in Abb. 5.3.2-3) als Nebenmaxima die Elektronensprünge unter Einbeziehung benachbarter Schwingungszustände auf (gestrichelte Linien in Abb. 5.3.2-3).

Für größere Moleküle ist das Franck-Condon-Prinzip sinngemäß anwendbar, wenn man sich eine Vorstellung darüber machen kann, welche Bindungen bei einem bestimmten Anregungszustand verändert werden. Oft sind dann nur ganz bestimmte Schwingungen geeignet, die Kompensation der veränderten Kernabstände zu erbringen, so daß nur diese im Spektrum auftreten.

Echte *Auswahlregeln* gelten zusätzlich für die *Elektronenübergänge*. Sie entsprechen den Kombinationsverboten der Atomspektroskopie, nach denen sich die Quantenzahl des Bahndrehimpulses nur um eine Einheit ändern darf, und daß Übergänge zwischen Niveaus verschiedener Multiplizität verboten sind. Für Moleküle gelten aber die Regeln nicht ohne

Einschränkungen, weil durch Molekülschwingungen die Symmetrie des Kerngerüsts vermindert werden kann. Daher werden Übergänge, die als reine Elektronenübergänge verboten sind, bei Kombination mit Schwingungsübergängen doch möglich. Diese *symmetrieverbotenen Übergänge* erreichen aber nur geringe Intensitäten. Im Benzenspektrum z. B. (Abb. 5.3.2-4) entspräche dem reinen Elektronenübergang eine Energiedifferenz von 38086 cm^{-1} ≙ 263 nm. Eine solche Absorption kommt nicht im Spektrum vor. Dafür sieht man eine große Anzahl stärkerer und schwächerer Absorptionen anschließend nach kürzeren Wellen. Jede einzelne gehört zu der gleichzeitig mit dem Elektronenübergang erfolgenden Schwingungsanregung. Von den $3n - 6 = 30$ Normalschwingungen des Benzens ist die totalsymmetrische Pulsation des Rings geeignet, die Bindungsverlängerung aller Ringbindungen auszugleichen, die infolge von deren Schwächung beim Übergang eines π-Elektrons auf ein höheres Niveau eintritt. Aus der Schwingungsstruktur entnimmt man demzufolge unter anderem Wellenzahldifferenzen von 923 cm^{-1}. Dieser Ringpulsation im angeregten Zustand entspricht die Wellenzahl von 990 cm^{-1} im elektronischen Grundzustand. (Die C–C-Kraftkonstante ist im elektronisch angeregten Zustand um 14% kleiner.) Die Schwingungsstruktur zeigt jedoch nicht nur die Wellenzahl 923 cm^{-1} der totalsymmetrischen Valenzschwingung. Die Anregung dieser Bewegung im oberen Zustand ändert noch nichts an der Symmetrie. Die Absorption kann nur dann eintreten, wenn gleichzeitig noch eine weitere, weniger symmetrische Schwingung angeregt ist. Das ist die unsymmetrische Ringschwingung von 520 cm^{-1}. Der Benzenring wird bei ihr zu einem länglichen Sechseck verzerrt.

Ein zweiter Absorptionsbereich folgt bei Benzen mit 45000 cm^{-1} bzw. 220 nm. Er ist zwar ebenfalls symmetrieverboten, tritt aber in ähnlicher Weise infolge der Mitwirkung von Schwingungen auf. Ein erlaubter Übergang liegt erst in dem mit gewöhnlichen Geräten nicht erfaßbaren Vakuum-UV mit dem Zentrum bei 55500 cm^{-1}. Die Absorption ist etwa 100mal stärker als die bei 45000 cm^{-1}.

Bei sehr großen Schichtdicken kann man eine sehr schwache vierte Absorption des Benzens bei 29400 cm^{-1} bzw. 340 nm feststellen. Sie ist einem Singulett-Triplett-Übergang zuzuschreiben und demnach *spinverboten*. Diese Absorption ist 10^5mal geringer als die des Bandensystems um 45000 cm^{-1}.

5.3.2.3.1. *Unscharfe Spektren.* Ein Vergleich des elektronischen Anregungszustands mit dem Grundzustand kann zuweilen zeigen, daß die Potentialkurve des oberen Zustands vom Grundzustand aus in dem Abschnitt erreicht wird (↑ Abb. 5.3.2-3), der oberhalb der Dissoziationsenergie liegt. Die Kurve beschreibt dann kein beständiges Molekül, sondern zwei durch starke äußere Kräfte enger als im Molekül aufeinandergeschobene Atome, die entweder ein Molekül bilden oder wieder auseinanderfliegen können. Dieses Gebilde hat keine Rotations- oder Schwingungszustände. Jeder Punkt auf der Potentialkurve bedeutet eine erlaubte Energie. Man erhält statt des Bandenspektrums

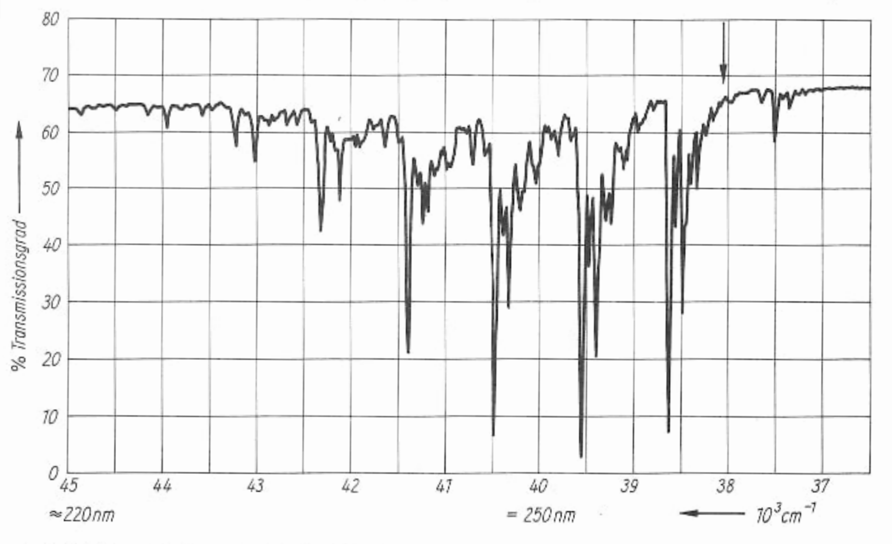


Abb. 5.3.2-4 Registrierung des UV-Absorptionsspektrums von Benzendampf

ein Kontinuum, an das sich noch eine Reihe von Banden nach längeren Wellen zu anschließen kann. Ein Beispiel ist das Absorptionsspektrum des Ioddampfs (Abb. 5.3.2-5). Wegen der kleinen Schwingungswellenzahl 213 cm^{-1} für dieses Molekül kann man sehr viele dicht benachbarte Schwingungsbanden beobachten.
In ähnlicher Weise können auch Zustände am Entstehen des Spektrums mitwirken, die überhaupt keinem stabilen Molekül entsprechen.



Abb. 5.3.2-5 Bandenspektrum mit Kontinuum, Absorption von Ioddampf

Für H_2 liegt oberhalb des Grundzustands ${}^1\Sigma_g^+$ ein Zustand ${}^3\Sigma_u^+$, d. h. mit ungepaarten Elektronen. Die berechenbare Potentialkurve zeigt kein Minimum; es gibt deshalb kein stabiles Molekül dieser Art und damit auch keine Rotations- und Schwingungsformen für einen solchen Zustand. Trotzdem tritt er spektroskopisch als Kontinuum in Erscheinung; ${}^3\Sigma_u^+$ bedeutet danach zwei mit gleichgerichteten Spins sehr eng benachbarte H-Atome, deren Abstand beliebig variieren kann, so daß die Energie auch kontinuierlich veränderlich ist. Bei noch höherer Energie gibt es einen Zustand ${}^3\Sigma_g^+$, zu dem die Potentialkurve eines Moleküls mit Energieminimum gehört. In diesen Zustand gelangt der Wasserstoff durch die Stöße der Elektronen in der Entladungslampe. Die Abstrahlung erfolgt durch den Übergang nach ${}^3\Sigma_u^+$ und liefert dabei das Kontinuum, das für die UV-Absorptions-Messungen bei kurzen Wellen als Lichtquelle benutzt wird. Von ${}^3\Sigma_g^+$ gibt es nur die zur Dissoziation führende Abstrahlung nach ${}^3\Sigma_u^+$, nicht aber zum Grundzustand ${}^1\Sigma_g^+$, denn bei leichten Atomen treten keine *spinverbotenen* Übergänge auf.

5.3.2.4. Absorptions- und Emissionsspektren, besondere Emissionen. Die Übergänge im Termschema können sowohl bei Absorption als auch bei Emission beobachtet werden. Bei normalen Temperaturen befinden sich alle Stoffe in ihren tiefsten Elektronen- und Schwingungszuständen, und Übergänge zu höheren Zuständen erfordern Energiezufuhr. Bei Absorption wird diese Energie einer den Stoff durchsetzenden Strahlung entzogen. Für Emissionsspektren dient der zu untersuchende Stoff selbst als Lichtquelle. In *thermischen Lichtquellen*, wie in Flammen oder Gasentladungen, werden die Moleküle des Stoffs durch Stöße von Elektronen, Atomen oder Ionen in angeregte Zustände überführt.
Bei verschiedenen Untersuchungen wirken die Moleküle als Strahlungsquellen, während sie selbst durch Strahlung angeregt werden oder angeregt wurden. Das erstere Phänomen ist die Lichtstreuung. Im Streulicht können nach Raman (↑ 5.3.4.2.) die Schwingungsfrequenzen des Moleküls gefunden werden. Es erfolgt eine teilweise Übertragung von Energie auf das Molekül, aber keine Absorption. Nach Absorption der anregenden Strahlung erfolgt die Emission bei *Fluoreszenz* ohne, bei *Phosphoreszenz* mit Zeitverzögerung. Emissionen dagegen, deren Ursache weder thermische Stöße noch Wellenstrahlen sind, werden als *Lumineszenz* bezeichnet.
Bei *Chemolumineszenz* emittieren die Moleküle Energie, die sie aus einer chemischen Reaktion ohne den Umweg über Wärmeentwicklung und nachfolgende thermische Anregung erhalten. Dazu gehört das Leuchten des weißen Phosphors im Dunklen infolge seiner langsamen Oxydation an der Luft, das mithin keine Phosphoreszenz ist. Weitere Beispiele bietet die Lichterzeugung bei Leuchtkäfern, -fischen und -bakterien, die auf chemischen Reaktionen in bestimmten Organen beruht.
Erfolgt die Anregung durch schnelle Elektronen oder hoch beschleunigte Ionen, so spricht man von *Katodolumineszenz.* Diese Erscheinung wird in Oszillographen- und Fernsehröhren, aber auch zum Nachweis von α-, β- und γ-Strahlen mittels Szintillationszählern in der Kernphysik benutzt.
Tribolumineszenz ist eine Lichterscheinung beim Zertrümmern von Kristallen. Das Sägen eines großen Zuckerblocks im gut verdunkelten Raum mit der Kreissäge ist von einer eindrucksvollen Lichterscheinung dieser Art begleitet.
Thermolumineszenz ist eine Sonderform der Phosphoreszenz. Sie besteht darin, daß eine bei tiefer Temperatur gespeicherte Anregungsenergie beim Erwärmen emittiert wird. Die letztgenannten Erscheinungen beruhen mehr auf Festkörper- als auf Moleküleigenschaften.

5.3.3. Spektroskopie im ultravioletten und sichtbaren Bereich

5.3.3.1. Absorptionsspektren

5.3.3.1.1. *Meßmethodik.* Als Lichtquellen mit kontinuierlichem Spektrum verwendet man meist Wasserstoff- oder Deuteriumgasentladungslampen (↑ 5.3.2.3.1.) für das kurzwellige UV von 60000 cm^{-1} bis 30000 cm^{-1} und Wolframdrahtglühlampen für den übrigen Bereich bis ins Ultrarot. Wie in der Atomspektroskopie ist *Quarz* ein geeignetes optisches Material für Prismen und Linsen; bei 50000 cm^{-1} bzw. 200 nm und größeren Wellenzahlen muß man zur Vakuumspektroskopie mit

Flußspatoptik übergehen, statt der Zerlegung der Strahlung durch Prismen ist jedoch die Anwendung von Gittern vorteilhafter, und zur optischen Abbildung dienen Spiegel. Die registrierenden Geräte empfangen die Strahlung mit einer Fotozelle oder mit einem Fotomultiplier. Die elektronische Verstärkung und Aufzeichnung des Spektrums durch ein Schreibwerk haben heute weitgehend die Spektrographen mit einer Fotoplatte verdrängt (↑ 4.3.2.1.).

Für Absorptionsmessungen werden die zu untersuchenden Gase und Flüssigkeiten in *Küvetten* gebracht, die aus UV-durchlässigem Quarz bestehen oder wenigstens Fenster aus diesem Material haben. Feste Stoffe werden meist in Lösung untersucht.

5.3.3.1.2. *Farbe und Absorption.* Die vollständige Deutung der Absorptionsspektren mit Hilfe der Potentialkurven der einzelnen Zustände und der Auswahlregeln ist nur für einfache Moleküle und nach umfangreichen Experimenten möglich. Andererseits ist ein wenigstens qualitatives Verständnis der Absorption auch komplizierter Stoffe notwendig, besonders wenn diese farbig sind oder gar als Farbstoffe verwendet werden.

Die Empfindung der Farbe entsteht durch das sichtbare Licht, das ein Stoff nicht absorbiert, also reflektiert. Weiß oder Farblosigkeit ist der Eindruck, den das Kontinuum des Sonnenlichts hervorruft. Obgleich das Zustandekommen der Farbempfindung kein physikalisches, sondern ein psychologisches Problem ist, wird im folgenden von Farbe auch zur Kennzeichnung der Absorptionseigenschaften gesprochen; beides hängt eng zusammen (Abb. 5.3.3-1). Im durchgehenden Licht erscheint ein Stoff danach in der Komplementärfarbe des absorbierten Lichts, Kupfersulfat z. B. blau, da es rotes und gelbes Licht absorbiert. Bei reiner Reflexion, z. B. an Metallen, entspricht der Farbeindruck dagegen direkt der absorbierten Farbe, da jeder Stoff die absorbierten Wellenlängen auch stark reflektiert. Goldaufdampfschichten sehen z. B. im durchgehenden Licht bläulich aus, massives Gold aber gelb.

Zur raschen Übersicht über die zu den Absorptionen gehörenden Komplementärfarben dient das Farbendreieck, in dem die drei reinen Grundfarben an den Ecken, die Mischfarben an den Seiten und die Komplementärfarben einander gegenüberstehen.

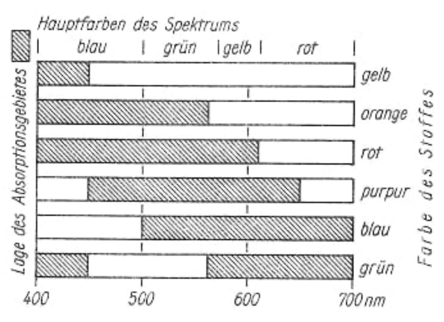


Abb. 5.3.3-1 Zusammenhang zwischen der Farbe und der Lage der Absorptionsbanden

Beschränkt man sich auf die Erklärung der Absorption im Sichtbaren, so braucht man für organische Stoffe nur die π-Elektronensysteme, für die anorganischen nur die d- und f-Elektronen zu betrachten, da die Anregung von σ-Elektronen zu Absorptionsbanden im fernen Ultraviolett führt.

5.3.3.1.3. *Farben organischer Verbindungen.* Die Absorption für die Anregung $\pi \rightarrow \pi^*$ eines π-Elektrons für Ethylen liegt bei 165 nm, also im Vakuum-UV. Die Konjugation mit einer weiteren Doppelbindung verschiebt sie nach 217 nm. Bei zunehmender Länge des konjugationsfähigen Systems gelangt man ins Sichtbare.

Abb. 5.3.3-2 Carotin, Möhrenfarbstoff, $\lambda_{abs} = 551$ nm

Abb. 5.3.3-3 Nonacarbocyanin; $\lambda_{abs} = 940$ nm (Infrarot); die aromatischen Ringe nehmen nicht an der Konjugation teil

Es gehören etwa 10 Doppelbindungen dazu, bis man zu roten oder orangeroten Stoffen gelangt, die allerdings sehr intensiv absorbieren, z. B. das Carotin (Abb. 5.3.3-2). Die Konjugation wird begünstigt, wenn das Farbstoffmolekül als Ion vorliegt. Dann genügen bereits weit kürzere Systeme. Dieser Art sind die *Polymethinfarbstoffe*, z. B. das Nonacarbocyanin (Abb. 5.3.3-3). Man kann in diesem Fall leicht Absorptionen bis hin zum Infraroten erhalten und diese Verbindungen als Sensibilisierungsfarbstoffe für die *Infrarotfotografie* verwenden.

Aromatische Verbindungen enthalten ebenfalls leicht anregbare π-Elektronen. Mit mindestens vier aromatischen Ringen sind diese Verbindungen farbig. Zu diesem Typ gehören die wegen ihrer Beständigkeit wertvollen *Indanthrenfarbstoffe*. In der älteren, empirischen Farbstofftheorie hat man für das System, das die relativ niedrigen Niveaus für die Elektronenübergänge bereitstellt – ehe man diese Vorstellungen kannte – den Begriff *Chromophor* [*chroma* griech., Farbe; *phoros* griech., tragend] geprägt. Chromophore in den obigen Beispielen sind die gesamten Konjugationsketten. Man gebraucht den Ausdruck auch für Gruppen wie C=O oder den Benzenring, obwohl da die Absorptionen im Ultravioletten liegen und keine Farbe sichtbar ist. Substituenten, die die Absorption von Chromophoren nach längeren Wellen verschieben, wurden *Auxochrome* [*auxe* griech., Vergrößerung] genannt.

5.3.3.1.4. *Anorganische Stoffe.* Eine Übersicht anhand des Periodensystems über das Auftreten von Farben bei anorganischen Stoffen läßt als größte Gruppe die mit den Ionen der Übergangselemente erkennen. Die Farbe beruht auf Übergängen der *d*-Elektronen bzw. bei der Gruppe der Lanthaniden und Actiniden auf denen der *f*-Elektronen. Man kann im Periodensystem weiter die Elemente zusammenfassen, die farbige Ionen oder farbige Moleküle unter geeigneten Umständen bilden können (Abb. 5.3.3-4). Diese Moleküle haben durchweg nichtbindende Elektronenpaare, und es muß sich bei den leichtanregbaren Elektronen um Mitglieder solcher Paare handeln. Verbinden sich verwandte Elemente dieser Art untereinander, so können wieder farbige Moleküle entstehen, z. B. P_4S_3, S_2Cl_2; aber PCl_3 und SF_6 sind farblos. Bor, Kohlenstoff, Silicium und besondere Modifikationen einiger anderer Elemente fallen durch ihre schwarze oder jedenfalls dunkle Farbe auf. Diese Stoffe sind Halbleiter, ihre schwarze Farbe hängt damit zusammen. Elemente aus der Nachbarschaft dieser Gruppe bilden miteinander entsprechende Verbindungen, z. B. Al_5P_3, Nitride, Sulfide.

Die Lichtabsorption der *d*-Elektronen kann nach der *Ligandenfeldtheorie* verstanden werden. Für die *d*-Elektronen gibt es fünf räumlich verschieden verteilte *d*-Orbitale (Abb. 5.3.3-5), die die fünffach entarteten Energiezustände der Elektronen im freien Atom darstellen. In Kristallgittern, in den Aquokomplexen, wie sie in wäßrigen Lösungen vorliegen, und in den sonstigen, meist auch komplexen Verbindungen trifft man die Ionen gewöhnlich in oktaedrischer oder tetraedrischer Koordination an. Obwohl das noch hochsymmetrische Anordnungen sind, ist damit die fünffache Entartung nicht verträglich. Es tritt eine Aufspaltung ein. Bei oktaedrischer Koordination sind das d_{z^2}- und das $d_{x^2-y^2}$-Orbital schlechter für die Besetzung mit Elektronen geeignet, weil diese Richtungen zu den Oktaederecken führen, welche die Liganden besetzen. Diese haben nichtbindende Elektronenpaare und sind oft sogar Anionen, so daß die Elektronen des Metallions abgestoßen werden. Umgekehrt ist es für tetraedrische Koordination. Dort gelangen die genannten Orbitale in die Lücken zwischen den Liganden, während die übrigen energetisch ungünstig liegen. Im ganzen bleibt sowohl für oktaedrische wie für tetraedrische Umgebung jeweils eine dreifache und eine doppelte Ent-

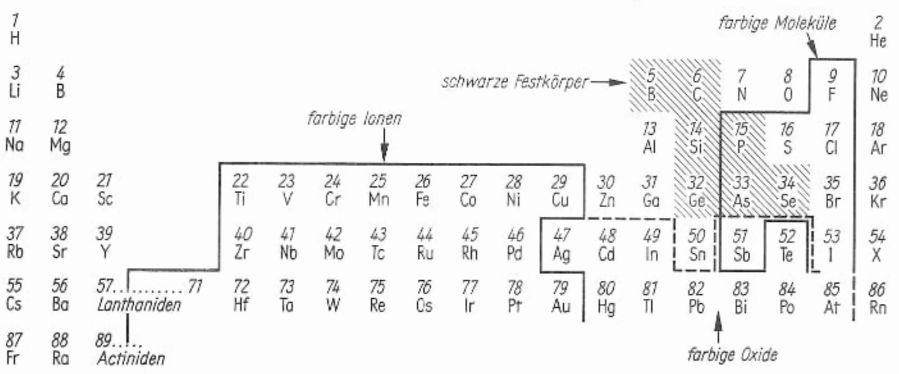


Abb. 5.3.3-4 Übersicht über anorganische farbige Verbindungen nach dem Periodensystem

artung bestehen. Diese Niveaus verschieben sich gegen die ursprüngliche mittlere Energie (Abb. 5.3.3-6). Sind nun die oberen Niveaus nicht voll besetzt, so sind Übergänge unter Absorption möglich. Die Größe der Aufspaltung hängt von der Stärke des Ligandenfelds ab.

Analyse identifiziert werden. Allerdings kommt dafür den Infrarotspektren (↑ 5.3.4.) eine größere Bedeutung zu, da viele Stoffe im Bereich von 200 nm bis 800 nm nicht absorbieren. Wird aber bei organischen Stoffen Absorption in diesem Bereich beobachtet, so müssen Gruppen

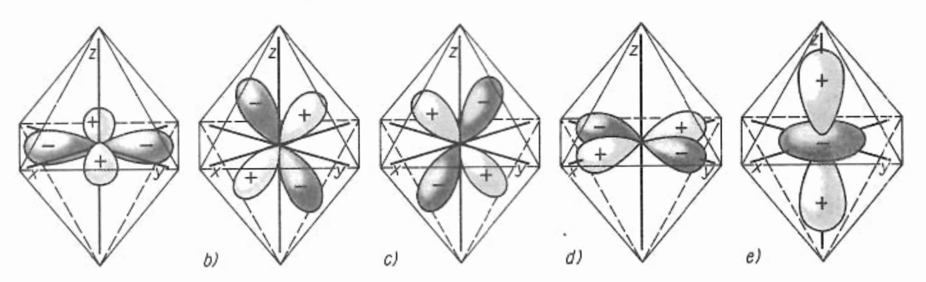


Abb. 5.3.3-5 Räumliche Erstreckung der *d*-Orbitale; bei oktaedrisch angeordneten Liganden a), b), c) günstig; *d*-Elektronen in den Lücken zwischen den Liganden

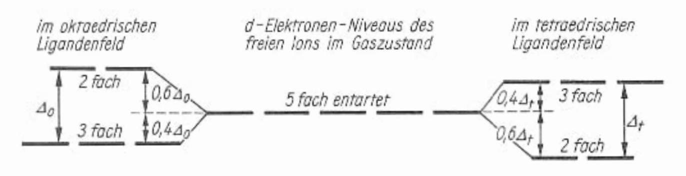


Abb. 5.3.3-6 *d*-Elektronenniveaus der Übergangselemente bei Komplexbildung oder Einbau in Kristallgitter

Die Einwirkung der Bindungspartner verstärkt sich in der Reihe $I^- < Br^- < Cl^- < F^- < OH^- < H_2O < NH_3 < C{\equiv}N^-$, d. h., die Absorptionsbanden werden nach kleineren Wellenlängen verschoben.

5.3.3.1.5. *Nicht auf Absorption beruhende Farben.* Einige Farberscheinungen werden nicht durch Absorption, sondern durch eine Reihe anderer physikalischer Effekte hervorgerufen. Die bevorzugte *Streuung* kurzer Lichtwellen erzeugt das *Blau des Himmels*, der sonst schwarz erscheinen müßte. *Interferenz* in dünnen Schichten liefert die farbigen Ringe eines *Ölflecks* auf nassem Untergrund. Durch *Brechung* entstehen die *Regenbogenfarben.* Auch die Farbe blauer Augen beruht auf dem Fehlen jeden Farbstoffs und allein auf Lichtstreuung in der Iris. Das Gelb blonden Haars wird durch sehr geringe Mengen des sonst braun und schwarz färbenden Melamins hervorgerufen. Laubfrösche erscheinen grün durch das Zusammenwirken eines gelben Farbstoffs mit dem durch Streuung in Körperzellen der Haut entstehenden Blau. Auch das schillernde Blau und Grün von Vogelfedern oder von Käfern wird nicht durch Farbstoffe, sondern durch einen besonderen Aufbau hervorgerufen.

5.3.3.1.6. *Analytische Anwendungen.* Viele Stoffe können an ihren UV-Absorptionsspektren erkannt und damit wie in der *chemischen*

mit π-Elektronen vorliegen. Besonders aromatische Moleküle und die zu diesen gehörigen heterozyklischen Verbindungen haben charakteristische Spektren mit vielen Einzelheiten, z. B. die Alkaloide, viele pharmazeutische Präparate und Produkte des lebenden Organismus. Die UV-spektroskopische Analyse wird deshalb in Pharmakologie, Medizin und Kriminalistik angewendet. Zu den ältesten spektroskopischen Methoden gehört der Nachweis der Kohlenoxidvergiftung anhand des veränderten Absorptionsspektrums des Bluts.

Die Absorptionsspektroskopie läßt sich auch als *quantitative Bestimmungsmethode* durchführen. Die Grundlage dafür ist das *Lambert-Beersche Gesetz* $E = \lg(\Phi_T/\Phi_0) = \varepsilon_\nu lc$ in dem E die Extinktion, Φ_0 die eingestrahlte Intensität, Φ_T die durchgelassene Intensität (T von Transmission), l die Schichtdicke, c die Konzentration der absorbierenden Substanz und ε_ν einen Proportionalitätsfaktor, den *Extinktionskoeffizient* bei der Wellenzahl $\tilde{\nu}$, bedeuten. Jede Absorptionsbande hat ihren eigenen maximalen Extinktionskoeffizienten ε_{max}.

Kennt man ε_ν für eine bestimmte Bande aus einer früheren Messung, ist l bekannt und die Transmission $T = \Phi_T/\Phi_0$ mit dem Spektralfotometer gemessen, so läßt sich die Konzentration c in mol l^{-1} berechnen. Gewöhnlich werden molare Extinktionskoeffizienten angegeben. Die Gültigkeit des Lambert-Beerschen

Gesetzes muß allerdings durch Eichmessungen mit bekannten Konzentrationen nachgeprüft werden, für die die reine Substanz zur Verfügung stehen muß, da konzentrationsabhängige Beeinflussungen der Extinktion durch Einwirkungen des Lösungsmittels oder durch andere gelöste Stoffe auf die absorbierenden Moleküle eintreten können und Assoziations- und Dissoziationsgleichgewichte grobe Veränderungen der Molekülkonzentration verursachen.

Viele Stoffe werden durch chemische Reaktionen erst in spektroskopisch nachweisbare Verbindungen übergeführt. Es ist dann vorteilhaft, ein im Sichtbaren absorbierendes, farbiges Reaktionsprodukt zu erzeugen, weil man dann schon mit bloßem Auge kontrollieren kann, ob die Reaktion oder die Messung störungsfrei erfolgt. Bei der Verlegung der Messung in das Sichtbare kann man auch Glasgefäße statt der kostspieligen Quarzküvetten verwenden.

5.3.3.2. Emissionsspektren

5.3.3.2.1. *Thermische Emission.* Da die meisten Moleküle bei hohen Temperaturen zerfallen, ist die thermische Anregung nur in einigen Spezialfällen wichtig; z. B. beruht der Nachweis von Erdalkalimetall in der Flamme des Bunsenbrenners mit Hilfe eines Handspektroskops nicht wie bei den Elementen Natrium und Kalium auf den Atomspektren der Metalle, sondern auf Banden. Die Quelle der für Calcium im Roten oder für Barium im Grünen gelegenen Emission, die auch zu Feuerwerken benutzt wird, ist ein dreiatomiges Molekül: Metall—O—H. Erst 1955 konnten durch leistungsfähige Spektrographen die Banden in Einzellinien aufgelöst und diese genau vermessen werden. In Flammen, die mit deuteriumhaltigem Acetylen brannten, traten gesetzmäßige Verschiebungen der Linien ein, aus denen auf die Anwesenheit von Wasserstoff im Molekül geschlossen werden konnte.

Praktisch wichtig sind Aussagen, die man aus Molekülemissionsspektren über die Verbrennungsvorgänge in Motoren, Gasturbinen und Raketentriebwerken erhält. Nach den spektroskopisch erkannten *Zwischenprodukten der Verbrennung* konnte man durch Zusätze Verbrennungsvorgänge beschleunigen oder auch abbremsen; allgemein bekannt als ein abbremsender Zusatz ist Bleitetraethyl als Antiklopfmittel in Benzinen. Auch die *Temperaturmessung* bei solchen Untersuchungen erfolgt spektroskopisch, da die Intensitätsverteilung der Rotationslinien eines zweiatomigen Molekülbruchstücks von der Temperatur in leicht auswertbarer Weise abhängt. Allerdings herrscht bei rasch verlaufenden Vorgängen überhaupt keine einheitliche Temperatur mehr, vielmehr haben die einzelnen Molekülarten je nach ihrer Beteiligung an der Reaktion und nach der dabei entwickelten Energie verschiedene „Temperaturen". Die Molekülemissionsspektren geben auch Auskunft über Vorgänge in den oberen Atmosphärenschichten, soweit sie von Lichtemission begleitet sind (Nordlicht), sowie über die Zusammensetzung von Kometen und der Atmosphäre von fremden Planeten (↑ 10.2.3.). Die besonderen Energieverhältnisse in den Flammen und in elektrischen Entladungen bedingen es, daß man vorzugsweise Spektren von Molekülen erhält, die bei normalen Temperaturen nicht existenzfähig sind und nach ihrer Entstehung nur eine kurze Lebensdauer haben. Man erfährt deshalb kaum etwas über die sonst in der Chemie untersuchten Verbindungen. Um solche Kenntnisse zu erlangen, werden in viel größerem Umfang Absorptionsspektren von Molekülen in Lösung untersucht.

5.3.3.2.2. *Fluoreszenz und Phosphoreszenz.* Licht, das von einem Molekül absorbiert wurde, kann auch wieder als Licht ausgestrahlt werden. Diese *Fluoreszenz* gibt Petroleum oder Motorenöl neben ihrer gelben bis braunen Farbe einen blauen Schimmer. Läßt man im verdunkelten Zimmer einen Sonnenstrahl durch die Flüssigkeit fallen, so hat das durchgegangene Licht gelbe Farbe, das blaue Fluoreszenzlicht aber tritt nach allen Richtungen wie aus einer Lichtquelle aus der Substanz aus. Für die Fluoreszenzstrahlung wird ein violetter Lichtanteil absorbiert; der Strahl, der durch das Gefäß mit dem Öl hindurchgegangen ist, kann in einer zweiten gleichen Probe keine Fluoreszenz mehr erregen. Die Erscheinung wurde nach den ersten Beobachtungen an Kristallen von natürlichem Flußspat (Fluorit) benannt. Sie ist dabei und bei einigen anderen anorganischen Stoffen eine Eigenschaft von *Fremdatomen im Kristallgitter* und geht deshalb beim Auflösen verloren. Demgegenüber zeigen organische Verbindungen Fluoreszenz als *Moleküleigenschaft*. Nur diese wird hier besprochen, obwohl die praktische Bedeutung der anorganischen Stoffe größer ist, da sie in Leuchtstoffröhren, zu Röntgen- und Fernsehbildschirmen und als Leuchtfarben Verwendung finden (↑ 5.3.3.1.).

Durch Einstrahlen einzelner durch Filter ausgesonderter Spektralbereiche einer Quecksilberdampflampe oder einer anderen Lichtquelle wird Fluoreszenzstrahlung in einer Meßküvette angeregt und in den Eintrittsspalt des Spektralfotometers geleitet. Wegen der geringen Intensität der Fluoreszenz sind die Messungen schwierig, und eine Verfälschung durch Reabsorption in der Probe muß rechnerisch korrigiert werden.

Das *Fluoreszenzspektrum* einer Substanz sieht dem Absorptionsspektrum ähnlich, da bei beiden Effekten dieselben Frequenzen zu

beobachten wären, wenn die Elektronenübergänge ohne gleichzeitige Änderungen von Schwingungszuständen erfolgten. Tatsächlich findet man, sofern die Schwingungsstruktur scharf ist, zunächst eine übereinstimmende Bande in beiden Spektren. Die weiteren Schwingungsbanden folgen aber bei der *Absorption nach kürzeren Wellen*, bei der *Fluoreszenz nach längeren Wellen* (Abb. 5.3.3-7). Bei der Absorption erfolgt der Übergang aus dem schwingungslosen Grundzustand in angeregte Zustände unter der Wirkung des Franck-Condon-Prinzips (↑ 5.3.2.3.), bei der Fluoreszenz aber vorzugsweise als Übergang aus dem Schwingungsgrundzustand des angeregten Moleküls in höhere Schwingungszustände des elektronischen

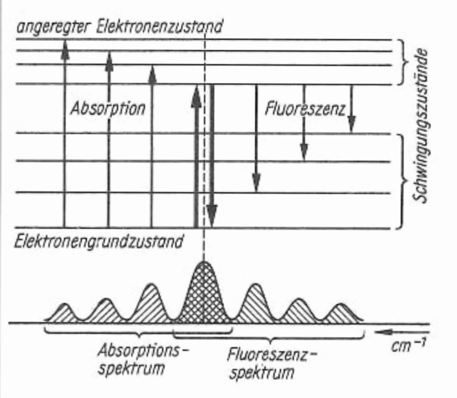


Abb. 5.3.3-7 Verwandtschaft zwischen Absorptions- und Fluoreszenzspektrum

Grundzustands. Infolgedessen sind auch die Abstände der Bandenmaxima verschieden. Weil im elektronisch angeregten Zustand die Kraftkonstanten und damit die Schwingungsfrequenzen kleiner sind als im Grundzustand, liegen im Absorptionsspektrum die Maxima näher beisammen als im Fluoreszenzspektrum. Daß sich dieses Spektrum stets mehr nach langen Wellen erstreckt als das Absorptionsspektrum, ist schon 1852 von STOKES erkannt worden. Fluoreszenzbanden, die ausnahmsweise kurzwelliger liegen, weil die absorbierenden Moleküle thermisch oder durch chemische Reaktionen bereits angeregt waren, bezeichnet man als *anti-Stokessche Banden*.

Bei der Betrachtung der Fluoreszenz als Umkehrung der Absorption ist es unerwartet, daß die große Mehrheit der chemischen Verbindungen keine Fluoreszenz zeigt. Selbst von den Stoffen, die π-Elektronen enthalten, fluoresziert nur ein Teil. Da viele aromatische Kohlenwasserstoffe recht intensive Fluoreszenz zeigen, während sie in ihren Derivaten kaum nachzuweisen ist, muß man annehmen, daß in der Regel der elektronisch angeregte, an sich fluoreszenzfähige Zustand durch eine rasche strahlungslose *Desaktivierung* an der Wiederausstrahlung gehindert wird. *Desaktivierungsprozesse* gibt es aber so viele, daß im Einzelfall schwer anzugeben ist, welcher der entscheidende war. Vielfach wird die absorbierte Strahlungsenergie zur Spaltung des Moleküls führen. Moleküle mit Doppelbindungen allerdings sind dagegen gesichert; denn wenn auch ein π-Elektron angeregt wird, so bleibt doch die σ-Bindung erhalten, und die Rückkehr in den Grundzustand ist möglich. Wenn aber Substituenten mit leichter spaltbaren Bindungen vorhanden sind, erfolgt an diesen die fotochemische Zerlegung.

Untersuchungen über Desaktivierungsvorgänge haben zu Ergebnissen über den Energietransport in chemischen Systemen geführt. Die elektronische Anregungsenergie wird ohne Stoß oder Auftreten von Strahlung an ein Nachbarmolekül weitergegeben, auch an irgendein andersartiges, wenn die Niveauunterschiede zufällig übereinstimmen.

Für den grundlegenden Lebensvorgang der Photosynthese der grünen Pflanzen ist derartige Energieübertragung wichtig. Chlorophyll *a* reagiert, optisch angeregt, direkt mit Wasser unter Bildung eines Radikalanions, dessen überschüssiges Elektron über viele Stufen schließlich als Reduktionsmittel auf Kohlendioxid wirkt. $[\text{Chl } a]^* + H_2O \rightarrow H^+ + \cdot OH + [\text{Chl } a]^-$; $4 \cdot OH \rightarrow 2\,H_2O + O_2$. Der Startprozeß ist aber nur möglich, wenn gleichzeitig Chlorophyll *b* anwesend ist und absorbiert.

Das äußere Kennzeichen der *Phosphoreszenz* ist die gegenüber der Fluoreszenz vergrößerte Zeitdauer zwischen Anregung und Wiederausstrahlung des absorbierten Lichtquants. Bei der hier allein betrachteten *Molekülphosphoreszenz* ist der Anregungszustand ein *metastabiler* Zustand, d. h. ein Zustand, von dem aus kein erlaubter Strahlungsübergang zum Grundzustand möglich ist. In den meisten Fällen handelt es sich dabei um einen *Triplettzustand*, bei dem zwei Elektronenspins parallel stehen.

5.3.3.3. Zeitauflösende Spektroskopie. Die Frage weniger oder mehr verzögerter Ausstrahlung nach Absorption, also nach Unterscheidung von Fluoreszenz und Phosphoreszenz, läßt das Bedürfnis nach einer Spektroskopie in Abhängigkeit von der Zeit entstehen. Nach langen Vorarbeiten hat sich durch die Entwicklung von Blitzlampen, schließlich aber durch die Schaffung der Laser dieses Feld zu einem wesentlichen Teilgebiet der modernen Spektroskopie entwickelt. Fotoblitzlampen haben eine Blitzdauer von 10^{-5} s, besondere Laser bis herab zu 10^{-12} s. Die Dauer des Anregungsimpulses muß kürzer sein als die Lebensdauer

des Anregungszustands, dessen Zeitverhalten untersucht werden soll. Man kann die Emission des angeregten Zustands messen. Rotierende Zylinder mit Blendenöffnungen dienten zunächst dazu, um zu bestimmten Zeiten nach der Anregung dem ausgestrahlten Licht den Weg zum Spektrographen freizugeben. Auch Absorptionsspektren kann man aufnehmen, wenn man dem Anregungsimpuls rasch eine lichtschwächere Beleuchtung mit einem kontinuierlichen Spektrum folgen läßt.

Laser dienen sowohl zur zeitlich definierten Anregung wie auch zum „Abfragen", d. h. zu einer der Anregung nachfolgenden Absorptionsmessung. Anregungs- und Abfrage-Impuls kann man z. B. mit einer optischen Verzögerungsstrecke zeitlich gegeneinander versetzen. Dabei wird der Lauf des Lichts in einem Spiegelsystem innerhalb des Laboratoriums benutzt, der immerhin 50 ns in Anspruch nehmen kann. Durch die Teilung des Anregungsimpulses an einem teildurchlässigen Spiegel wird ein Strahl abgezweigt, der um einen genauen Zeitbetrag später auf die Substanz trifft als der Anregungsimpuls. Schaltungstechnisch wird mit Hilfe einer 2-Kanal-Tor-Elektronik gewährleistet, daß nur für die Absorptionsmessung der Empfänger eingeschaltet ist. Es kann aus der Absorptionsintensität festgestellt werden, wieviele Moleküle (relativ) inzwischen wieder in den Grundzustand zurückgekehrt sind. Durch Verwendung verschiedener Laser mit unterschiedlichen Frequenzen kann man die Methodik beliebig ausbauen, z. B. auch durch Messung mit einer Frequenz entsprechend der Energie des Anregungszustands feststellen, wieviel Moleküle sich zu bestimmten Zeiten in diesem befinden.

5.3.4. Infrarot- und Ramanspektroskopie

5.3.4.1. Infrarotspektralfotometer. Im infraroten Spektralgebiet findet man Absorptionen infolge von Schwingungsübergängen (einschließlich meist unaufgelöster Rotationsstruktur); bei langen Wellen können sich Absorptionen der reinen Molekülrotation anschließen. Wie für die Untersuchung der Absorptionsspektren im Ultravioletten und Sichtbaren (↑ 5.3.3.1.) muß ein Infrarotspektralfotometer im einfachsten Fall aus einer Strahlungsquelle mit kontinuierlichem Spektrum, aus einem Monochromator mit Gitter oder Prisma und aus einem geeigneten Empfänger bestehen.

5.3.4.1.1. *Infrarotstrahlungsquellen.* Obgleich jeder glühende Stoff ein kontinuierliches Spektrum emittiert, ist nicht jeder als Strahlungsquelle geeignet, da die Intensität der Emission in verschiedenen Wellenlängenbereichen verschieden ist. Aus an sich diskreten Gitterschwingungen entsteht ein kontinuierliches Spektrum nur, weil die Linien bei hoher Temperatur durch Kopplungseffekte enorm verbreitert sind. Nur ein *schwarzer Strahler* oder *schwarzer Körper* emittiert in allen Spektralbereichen gleichmäßig gut, da nach KIRCHHOFF das Emissions- dem Absorptionsvermögen gleich ist; kein Stoff erreicht aber dieses theoretische Emissionsvermögen.

In der Energieverteilung des schwarzen Körpers entfällt nur eine geringe Intensität auf das Gebiet von 2,5 bis 25 µm. Die Infrarotspektroskopie kann deshalb nur mit geringen Intensitäten experimentieren. Da die Steigerung der Strahlertemperatur das Emissionsmaximum weiter zum Sichtbaren hin verschiebt, muß man Temperaturen von 1400 bis 2000 °C anstreben,

Tab. 5.3.4-1 Strahlungsquellen für die Infrarotspektroskopie

Name	Material	Leitvermögen	Leistungsaufnahme	Emissionsvermögen des schwarzen Körpers	Temperatur
Silitstab [engl. Globar]	Siliciumcarbid SiC	Halbleiter	300 bis 500 W	75 bis 80% von 2,5 bis 25 µm	max. 1350 °C
Nernststift	ZrO_2 85% Y_2O_2 u. a.	ionenleitend ab 800 °C	50 bis 100 W	75 bis 80% von 5 bis 15 µm dann Rückgang auf 50%	max. 1900 °C
Metallwendel (Metallfadenlampen)	z. B. Cr-Ni oder W	metallischer Leiter	1000 W	durch Oberflächenoxide oder Carbide 90% erreichbar, bis 15 µm brauchbar	1000 bis 1500 °C
Keramik oder Sinteroxidrohr	z. B. Al_2O_3	Isolator geheizt durch metall. Leiter (Pt, Rh)	z. B. 30 W	bis 25 µm brauchbar	1200 °C

um nur eine geringe Verbesserung der Emission im Bereich der Molekülschwingungen zu erhalten. Nur wenige Stoffe können aber so hoch erhitzt werden (Tab. 5.3.4-1).

Die in kürzeren Wellenlängen emittierte Hauptmenge der Energie thermischer Strahler geht nicht nur verloren, sondern behindert auch die Messungen bei langen Wellen, weil minimale, bei dem großen Überschuß aber spürbare Anteile des kurzwelligen Spektrums als *Streustrahlung* in den optischen Geräten im untersuchten langwelligen Spektrum auftreten. Deshalb sind Versuche interessant, eine monochromatische Strahlungsquelle zu entwickeln, deren Wellenlänge sich über einen größeren Bereich kontinuierlich verstellen läßt. Periodisch rasch bewegte Elektronen können dafür geeignete Wellen aussenden, was man als *Smith-Purcell-Effekt* bezeichnet. Durchstimmbare Strahlungsquellen, allerdings für nur kleine Spektralbereiche ($\approx 100\ cm^{-1}$), stellen jeweils die Halbleiterlaserdioden dar. Diese Kristallspaltstücke von Kubikmillimeter Größe senden gerichtete infrarote Strahlung aus, die mit einer außerordentlich geringen Bandbreite erhalten werden kann ($0{,}0001\ cm^{-1}$ im Einzelmodenbetrieb) und damit eine höchstauflösende Spektroskopie gestattet. Das dabei verwendete Beugungsgitter dient nur zur Selektion der einzelnen Lasermoden.

5.3.4.1.2. *Monochromator.* Für das Infrarot verwendet man keine Spektrographen, die das gesamte Spektrum gleichzeitig zeigen. Da man es nicht fotografiert, genügen Monochromatoren.

Für die optischen Bauteile müssen *infrarotdurchlässige Materialien* verwendet werden. Glas ist nur bis 2 μm brauchbar. *Natürliche Steinsalzkristalle*, die sehr selten optische Qualität haben, waren lange Zeit das einzige Hilfsmittel. Heute werden Einkristalle nicht nur für piezoelektrische und Halbleitermaterialien gezüchtet, sondern auch für die Infrarotspektroskopie. Wegen der unterschiedlichen Anforderungen an die chemische oder an die atmosphärische Widerstandsfähigkeit und weil man kein Prisma herstellen kann, das das ganze Gebiet von 2 bis 40 μm gleichzeitig auflöst, muß eine ganze Reihe solcher Stoffe verfügbar sein. Die Dispersion ist nur dann groß, wenn in benachbarten Spektralbereichen Absorptionsgebiete liegen. Die Änderung des Brechungsindex mit der Wellenlänge kann als beginnende, die Absorption als vollständige Wechselwirkung der Strahlung mit der Materie aufgefaßt werden. Glas ist für Prismen im sichtbaren Bereich nur geeignet, weil es im Ultraviolett undurchlässig wird (Tab. 5.3.4-2).

Prismen allein werden in kleineren Geräten für einfache Aufgaben verwendet. Für höhere Auflösung dienen Prismen nur zu einer Vorzerlegung, die Hauptzerlegung erfolgt durch Gitter. Dabei bleibt es aber eine Besonderheit der Infrarotspektroskopie mit thermischen Strahlern, daß sie mit energiebegrenzter Auflösung arbeitet. Man muß die Spalte gewöhnlich weiter öffnen, als es für die Ausnutzung des Prismas oder Gitters zweckmäßig ist, um noch genügend Energie für die Messung zu haben. So erreichen nur kostspielige Geräte $0{,}1\ cm^{-1}$ Auflösung über den Bereich von 250 bis $4000\ cm^{-1}$.

Tab. 5.3.4-2 Fenster- und Prismenmaterialien für die Infrarotspektroskopie; *D* Dispersionsgebiet in cm^{-1} für Stoffe, die für Prismen geeignet sind; *G* langwellige Grenze in cm^{-1} der Durchlässigkeit

Material	*G*	*D*
Glas	5000	–
Quarz	2500	–
Glimmer	2000	–
LiF	1500	5000 bis 1885
CaF_2	1200	1885 bis 1175
As_2S_3-Glas	900	–
Irtran-2, ZnSe polykristallin	750	–
NaCl	600	1175 bis 650
KBr	350	650 bis 400
AgCl	350	–
KRS-5, Tl [Br, I]	250	500 bis 285
CsI	200	400 bis 250

Zur Herstellung von Küvetten sind noch viele andere Stoffe, z. B. Diamant, Silicium, aber auch Plastfolien (Polyethylen, Teflon) verwendet worden.

5.3.4.1.3. *Interferometrische Infrarotspektroskopie.* Die Probleme der geringen Strahlerenergie haben in der Infrarotspektroskopie zu Geräten geführt, bei denen durch ein Interferometer ein Interferenzbild erzeugt und registriert und aus diesem das Spektrum errechnet wird. Beide Darstellungen enthalten im Grunde die gleiche Information, einerseits die Abhängigkeit der Brechungszahl von der Wellenlänge und andererseits die Wellenlängenabhängigkeit des Absorptionsvermögens. Durch Fouriertransformation kann die eine Darstellung in die andere umgerechnet werden (↑ 2.3.1.). Die Aufnahme eines Interferogramms ist meßtechnisch günstiger. Das Interferometer (vgl. zum Aufbau mit Michelsonversuch zur Relativitätstheorie, 1.2.2.) hat nicht wie der Monochromator Eintritts- und Austrittsspalt, sondern Kreisblenden, so daß ein Vielfaches an Strahlung hindurchgeht. Da keine Dispersion erfolgt, werden alle Wellenlängen während des Versuchs gleichzeitig gemessen. Die Schwierigkeiten

lagen bei den Ansprüchen an die mechanische Präzision einiger Interferometerbauteile und an eine genaue Intensitätsmessung über fünf Zehnerpotenzen. Entwicklungen für die Kosmosforschung kommen heute der Infrarotspektroskopie zugute. Diese Geräte messen bei beliebig hoher Auflösung (die vom Weg des beweglichen Interferometerspiegels abhängt) Spektren geringer Stoffmengen, z. B. in Oberflächenschichten, und dann auch vorteilhaft als Emissionsspektren, wenn sich die Proben z. B. nur 100 grd über der Umgebungstemperatur befinden.

5.3.4.1.4. *Strahlungsempfänger.* In *thermischen Empfängern* wird die einfallende Strahlungsenergie in Wärme umgewandelt und eine Wirkung der damit erzeugten Temperaturerhöhung gemessen; in *Quantenempfängern* löst das Strahlungsquant $h\nu$ direkt einen Effekt aus, z. B. einen *lichtelektrischen Effekt*, der im Ultravioletten oder Sichtbaren leicht, an Materialien wie Bleisulfid oder besser Indiumantimonid aber auch im Ultraroten erzeugt werden kann. Diese Stoffe haben aber nur in einem bestimmten Bereich eine hohe Empfindlichkeit, sie sind *selektiv*, während die thermischen Empfänger *universell* und prinzipiell vom Ultraviolett bis zum äußersten Ultrarot verwendbar sind. Lange bekannt ist das *Bolometer*, in dem durch die Temperaturerhöhung der elektrische Widerstand einer angestrahlten Folie verändert und gemessen wird. Günstigere Eigenschaften haben *Kristallbolometer*. Die Strahlung wird in einem nicht zu kleinen pyroelektrischen Kristall, gewöhnlich von Triglycinsulfat (TGS) absorbiert. Wegen der kurzen Ansprechzeit (≙ Zeitkonstante) sind TGS-Empfänger für die Fouriertransformations-IR-Spektroskopie unentbehrlich. Beim *Thermoelement* fließt ohne äußere Spannung ein Strom, wenn eine Temperaturdifferenz zwischen den beiden Verbindungsstellen verschiedener ausgewählter Metalle entsteht. In pneumatischen Empfängern wird die durch die Erwärmung bedingte Druck- oder Volumenzunahme eines Gases in einen elektrischen meßbaren Effekt umgewandelt.

Durch Wechselstromverstärker und Bandfilter kann das Rauschen der Empfänger von der Verstärkung ausgeschlossen werden, wenn man *Wechsellicht* verwendet; man erhält es, indem eine rotierende Scheibe mit ausgeschnittenen Sektoren den Strahl mit einer bestimmten Frequenz unterbricht. Führt man die Sektoren der Scheibe als Spiegel aus, so kann abwechselnd Licht aus dem Meßstrahl und aus einem Vergleichsstrahl durch den Monochromator zum Empfänger geschickt werden.

5.3.4.2. Ramaneffekt. Er wurde nach Voraussagen von SMEKAL u. a. und durch Untersuchungen von RAMAN sowie von LANDSBERG und MANDELSTAM gefunden und zeigt ebenfalls die Schwingungen und Rotationen der Moleküle und zwar in einem Experiment der Lichtstreuung. Durch Einstrahlen monochromatischen Lichts bestimmter Frequenz ν_0 wird die Elektronenhülle eines Moleküls zu erzwungenen Schwingungen veranlaßt und wirkt als Wellensender. Das ausgesandte Streuspektrum enthält wegen der Schwingungen der Atomkerne des Moleküls (und wegen der Rotationen bei Untersuchungen im Gaszustand) neben Linien der Frequenz ν_0 des anregenden monochromatischen Lichts noch *Ramanbanden* der Frequenz $\nu = \nu_0 (\pm) \nu_i$, wenn ν_i die Schwingungs- oder Rotationsfrequenz des streuenden Moleküls ist. Von der Fluoreszenz unterscheidet sich der Vorgang darin, daß die eingestrahlte Energie noch nicht zur elektronischen Anregung ausreicht. Frequenzerhöhung oder das Auftreten antistokesscher Linien ist möglich, wenn sich die Moleküle in einem entsprechend angeregten Schwingungszustand bereits befinden.

Die Untersuchung des Ramaneffekts ist 35 Jahre lang hauptsächlich durch Einstrahlen intensiver Spektrallinien des Quecksilbers (besonders der blauen Linie 435,8 nm) vorgenommen worden. Erst mit der Entwicklung der Laser ist die optimale Strahlungsquelle für die Ramanspektroskopie geschaffen worden, die eine bis 10^{10}mal intensivere Einstrahlung erlaubt. Die Anwendung der Laser hat es auch ermöglicht, zwei Gebiete einer nichtklassischen Ramanspektroskopie auszubauen bzw. überhaupt erst zu entdecken: Resonanzramaneffekt (↑ 5.3.4.2.2.) und nichtlineare Ramaneffekte (↑ 5.3.4.2.3.).

5.3.4.2.1. *Untersuchungsmethodik des klassischen Ramaneffekts.* Die klassische Raman-

Tab. 5.3.4-3 Laser der Ramanspektroskopie, λ Wellenlänge, P Leistung;
*) YAG = Yttrium-Aluminium-Granat
**) (C_2H_5) HN … O … $NH^+C_2H_5]Cl^-$ … C … $C_6H_5—COO—C_2H_2$

Typ	Medium	λ in nm	P in W
Kristall (gepulst)	Rubin Nd-YAG*)	694,3 530	1 100
Gas (kontinuierlich)	He–Ne Ar^+ Kr^+	632,8 514,5 488,0 647,1	bis 0,18 1 0,6
Farbstofflösungen	z. B. Rhodamin 6G**)	530 bis 680 regelbar	≈0,3

spektroskopie wird heute nur noch mit Lasern als Strahlungsquellen bearbeitet; Tab. 5.3.4-3 enthält Angaben über gebräuchliche Typen.
Verschiedene Anregungswellenlängen sind erforderlich, da man außerhalb von Absorptionsgebieten der Substanzen einstrahlen muß, um Absorption und Fluoreszenz zu vermeiden. Unter diesen Bedingungen ist der Fluß der Ramanstreustrahlung

$$(5)\quad \Phi_{\mathrm{Ra}} \sim \frac{(\nu_0 - \nu)^4}{1 - \exp(-h\nu/k_B T)} \left(\frac{\partial\alpha}{\partial q_i}\right)^2 \Phi_0,$$

wobei k_B Boltzmannkonstante und T Temperatur bedeuten. Die Ramanstreustrahlung hängt also von der Wellenzahl der Erregerstrahlung stark ab; die Verwendung kurzwelligen Lichts ist vorteilhaft. $\partial\alpha/\partial q_i$ ist die Polarisierbarkeitsänderung des Moleküls während einer Schwingung, ein Maß für die mehr oder weniger leichte Verformbarkeit der Elektronenhülle im Feld der einfallenden Strahlung. Benzen war eine günstige Substanz für die Entdeckung des Ramaneffekts, weil die π-Elektronen leicht beweglich sind.
Nur ein geringer Teil der eingestrahlten Energie wird in Ramanstreuung umgesetzt. Der Streustrahlungsfluß von einem Einzelmolekül ist $\Phi \approx 10^{-30}\Phi_0$ in günstigen Fällen. So waren ältere Versuchsanordnungen für Gase darauf gerichtet, Strahlung von 10 l Probe in den Spektrographen zu sammeln. Mit Laseranregung erreicht man bessere Resultate mit Bruchteilen von Millilitermengen, für die dann $\Phi \approx 10^{-9}\Phi_0$ gilt. Daneben tritt um mindestens eine Größenordnung intensivere molekulare Streustrahlung ungeänderter Wellenzahl auf, die *Rayleighstreuung*, und die auf optischen Inhomogenitäten (Staubteilchen, Luftblasen, örtlich verschiedenem Brechungsindex) beruhende *Tyndallstreuung*.
Um die lichtschwache Ramanstreuung neben der rund 10^9fach stärkeren Primärstrahlung beobachten zu können, untersucht man das Streulicht senkrecht zur Primärstrahlungsrichtung. Rayleigh- und Tyndallstreuung, die ebenfalls diese Hauptrichtung haben, aber die unveränderte Wellenzahl $\tilde{\nu}_0$, werden durch Monochromatoren besonderer Streulichtfreiheit abgetrennt. Dazu sind Doppel- und sogar 3fach-Monochromatoren gebaut worden. Mit *holographisch erzeugten Gittern*, die weit größere Präzision als die mechanisch hergestellten haben, erreicht man noch bessere Ergebnisse.
Als Empfänger dienen Sekundärelektronenvervielfacher, die nach dem Gesichtspunkt großer Empfindlichkeit bei kleinem Dunkelstrom aus der Produktion herausgesucht sind und zur Herabsetzung des Dunkelstroms noch gekühlt werden können.

5.3.4.2.2. *Resonanzramaneffekt.* Läßt man die Bedingung beiseite, daß die Einstrahlung der Erregerfrequenz in größerer Entfernung von einem Absorptionsgebiet erfolgen soll, und strahlt mit einer Wellenzahl ein, die nur wenig höher als die einer Absorptionsbande ist, so wird der Ramaneffekt bedeutend verstärkt, ist aber nur schwer zu beobachten, da gleichzeitig auch Reabsorption erfolgt. Man muß deshalb sehr verdünnte Lösungen anwenden. Shorygin wies den Effekt an gelb erscheinenden Nitroverbindungen bei Anregung mit der blauen Quecksilberspektrallinie von 435,8 nm in einer Verdünnung 1 : 10000 nach. Es handelte sich dabei und bei den meisten Anwendungen um den *Präresonanzramaneffekt*. Bei Einstrahlung direkt in den Bereich einer Absorptionsbande, dem *rigorosen Resonanzramaneffekt*, beobachtet man auch noch Reihen höherer Oberschwingungen von einigen Grundschwingungen.
Der Resonanzramaneffekt hat Anwendungen gefunden für Untersuchungen an den stark verdünnten wäßrigen Lösungen biochemischer Systeme und zu Untersuchungen über größere Entfernungen. Mit einer passenden Laserfrequenz konnte z. B. vom Flugzeug aus der Iodgehalt des Meerwassers durch Registrierung des Resonanzramansignals verfolgt werden. Die Farbstofflaser mit einstellbaren Wellenlängen sind wichtig für dieses Arbeitsgebiet.

5.3.4.2.3. *Nichtlineare Ramaneffekte.* Bei der Einstrahlung von weit höheren Laserleistungen, die allerdings von kürzester Dauer sein können und die mit *Riesenimpulslasern* erreicht werden, gibt es eine ganze Reihe besonderer Effekte, die darauf beruhen, daß die in 5.3.(5) benutzte Polarisierbarkeit keine Konstante mehr ist, sondern eine Funktion der elektrischen Feldstärke (in der Lichtwelle), die daher mit höheren Potenzen in die Beschreibung der Wirkungen eingeht. Erste dieser Entdeckungen war der *stimulierte Ramaneffekt*: Bei Riesenimpulseinstrahlung tritt Ramanstrahlung $\nu = \nu_0 - \nu_i$ der stärksten Ramanlinie der Substanz in Laserform, gerichtet aus der Substanz aus. Der Strahl enthält eine Folge der Oberschwingungen dieser Frequenz $\nu = \nu_0 - n\nu_i$, aber nichts von sonstigen Ramanfrequenzen. Es treten weiter auch Frequenzen $\nu = \nu_0 + \nu_i$ auf (↑ 5.3.4.2.), ohne daß angeregte Moleküle vorausgesetzt werden können, und zwar in besonderen Strahlen, die kleine Winkel gegenüber dem Strahl mit ν_0 bilden und weniger scharfe Frequenzen haben: $\nu = \nu_0 + \nu_i \pm \delta\nu$. Diese anti-Stokes-Strahlung des stimulierten Ramaneffekts wird von der gleichen Substanz absorbiert, andere Substanzen mit Frequenzen ν_{i2} absorbieren aber bei $\nu_0 + \nu_{i2}$, falls diese Frequenz in das mit $\pm\delta\nu$ angedeutete Kontinuum fällt, man spricht vom *inversen Ramaneffekt*.
Bei der *kohärenten anti-Stokes-Raman-Spek-*

troskopie (CARS) werden gleichzeitig die Frequenz ν_1 eines konstanten und ν_2 eines abstimmbaren Lasers in eine Substanz eingestrahlt (gepulste Laser). Ähnlich dem stimulierten Ramaneffekt entsteht ein Antistokes-Strahl mit $\nu_{as} = 2\nu_1 - \nu_2$, wenn $\nu_1 - \nu_2 = \nu_i$ einer Molekülschwingung ist. So kann man $\nu_{as} = \nu_1 + \nu_i$, also das anti-Stokes-Raman-Spektrum einer Substanz registrieren, und zwar unter günstigen Energieverhältnissen, weil es in Laserform aus der Substanz austritt. Damit erfolgt seine Aufnahme auch ungestört von Fluoreszenz oder anderen Substanzemissionen. Man hat Ramanspektren von Flammen in dieser Weise aufgenommen.

5.3.4.3. Anwendungen der Infrarot- und Ramanspektroskopie

5.3.4.3.1. *Identifizierung und Analyse.* Während es bei der Kennzeichnung chemischer Verbindungen durch den Schmelzpunkt oder den Brechungsindex noch an die hundert Stoffe mit den gleichen Konstanten geben kann, ist das *Infrarotspektrum* trotz starker Ähnlichkeiten zwischen den Spektren verwandter Substanzen völlig spezifisch und dient deshalb zur Kennzeichnung der Identität und der Reinheit von Substanzen. Eine Gegenüberstellung von Ultraviolett- und Infrarotspektrum (Abb. 5.3.4-4, 5.3.4-5) zeigt, daß das Infrarotspektrum durch die große Anzahl von Banden verschiedener Intensität eine Substanz viel schärfer bezeichnet.

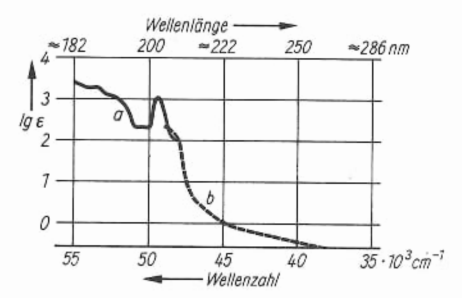


Abb. 5.3.4-4 UV-Absorptionsspektrum von Cyclohexen; a) gasförmig, b) gelöst in Cyclohexan

Selbstverständlich läßt sich auch die quantitative Analyse von Stoffen in Mischungen anhand der Absorptionsintensität einzelner Banden mit Hilfe von Eichkurven und des Lambert-Beerschen Gesetzes so vornehmen, wie sie bei der Anwendung der UV-Spektren beschrieben wurde. Nur sind in der Infrarotspektroskopie noch weitere Umstände bei der Eichung der Verfahren zu beachten; z. B. verfälschen die aus meßtechnischen Gründen großen Spaltbreiten der Infrarotspektrometer die Bandenformen. Eine gewisse Einschränkung liegt auch darin, daß nicht viel mehr als drei Komponenten zugegen sein dürfen, weil die sich überlagernden Spektren dann nicht mehr getrennt werden können (↑ 5.3.4.4.).

5.3.4.3.2. *Charakteristische Gruppenschwingungen.* Die Schwingungsspektren enthalten kodiert die *Konstitution* der chemischen Verbindung. Für viele Atomgruppierungen gibt es charakteristische Frequenzen, die als Infrarotabsorptionen oder Ramanbanden an bestimmter Stelle im Spektrum auftreten. Möglich sind solche charakteristischen Schwingungen allerdings nur, wenn sich durch bedeutende Masseunterschiede oder durch einen größeren Unterschied einer Kraftkonstanten von den übrigen eine von den restlichen Schwingungen unabhängige Bewegung innerhalb einer Gruppe ausbilden kann. Grundsätzlich schwingen alle Atome eines Moleküls gemeinsam. Es gibt aber unter den angeführten Voraussetzungen Bewegungsformen, bei denen sich die Schwingungsenergie in einer Bindung oder in der Schwingung zwischen wenigen Massen so weit konzentriert, daß die gleichzeitigen Bewegungen

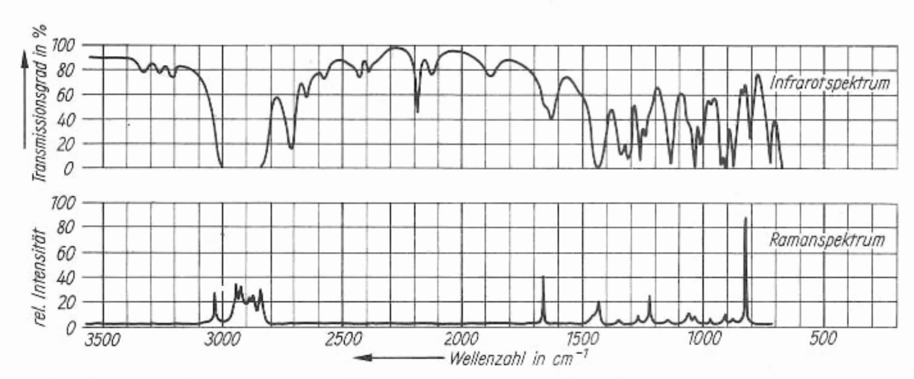


Abb. 5.3.4-5 Infrarotspektrum und Ramanspektrum von Cyclohexen

der übrigen Atome dagegen vernachlässigt werden können. Wasserstoff mit seiner geringen Atommasse führt immer zur Ausbildung derartiger charakteristischer Schwingungen. Die *X*-H-Valenzschwingungen der einzelnen Elemente *X* zeigen eine Systematik, die dem Periodensystem folgt (Tab. 5.3.4-6). Nach $\nu = \frac{1}{2\pi}\sqrt{\frac{f}{m}}$ (↑ 5.3.1.2.) ist daraus zu schließen, daß die Bindung im Fluorwasserstoff die festeste ist, die Wasserstoff eingehen kann.

Die großen angegebenen Bereiche bei einzelnen *X*-H-Wellenzahlen bedeuten keinesfalls eine Unsicherheit unserer Kenntnisse; vielmehr gibt es gesetzmäßige Unterschiede. In erster Linie ist – und das ist bei Kohlenstoff besonders festzustellen – die Möglichkeit verschiedener Bindungshybridisierung am *X*-Atom maßgebend (Tab. 5.3.4-7). *Je höher der s-Orbital-Anteil in der C-H-Bindung, desto höher die Wellenzahl.* Es kommen mehrere C-H-Schwingungen zur Beobachtung (entweder in Infrarotabsorption oder im Ramaneffekt), wenn es mehrere C-H-Valenzen wie in den angeführten Beispielen gibt, zwischen deren Schwingungen es zu Kopplungen mit gegenseitiger Beeinflussung der Frequenzen kommt [↑ 5.3.(2)].

Weitere Einwirkungen bringen induktive und mesomere Effekte mit sich. *Negativierung des C-Atoms führt zu erhöhten, Positivierung zu erniedrigten C-H-Schwingungen* (Tab. 5.3.4-8).

Da die Valenzschwingungen durch das Auftreten von Oberschwingungen niedriger liegender Grundschwingungen in diesem Bereich gewöhnlich gestört sind, beobachtet man diese Unterschiede besser an den entsprechenden symmetrischen Deformationsschwingungen der Methylgruppe. Diese andersartige Bewegung der gleichen Gruppe muß selbstverständlich die gleichen Gesetzmäßigkeiten zeigen (Tab. 5.3.4-9).

Ein Beispiel für einen mesomeren Effekt der Frequenzbeeinflussung ist die besonders niedrige Lage der aldehydischen C-H-Valenzschwingung: 2820 cm^{-1}. In diesem Fall erfolgt die Positivierung des Kohlenstoffs durch die Mesomerie $\rangle C{=}O\rangle \leftrightarrow \rangle \overset{\oplus}{C}-\overline{\underline{O}}|^{\ominus}$

Tab. 5.3.4-6 Systematik der *X*-H-Valenzschwingungen der Elemente *X* nach dem Periodensystem; Angaben in Wellenzahlen je Zentimeter

B	C	N	O	F
2200 bis 2640	2800 bis 3374	2900 bis 3400	3200 bis 3700	3935
Al	Si	P	S	Cl
1670 bis 1810	2250 bis 2100	2440 bis 2275	2600 bis 2300	2885
	Ge	As	Se	Br
	2010 bis 2160	≈2100	2250 bis 2350	2558
	Sn	Sb	Te	I
	1800 bis 1900	≈1890		2233

Tab. 5.3.4-7 Unterschiedliche C-H-Valenzschwingungen infolge unterschiedlicher Hybridisierung der C-Atome

Verbindung	Formel	Wellenzahl in cm^{-1}
Acethylen (=Ethin)	$CH{\equiv}CH$	3287, 3374
Ethylen (=Ethen)	$CH_2{=}CH_2$	2989 bis 3272
Cyclopropan	C_3H_6	3024 bis 3103
Benzen	C_6H_6	3047 bis 3099
Methan	CH_4	2914, 3020

Tab. 5.3.4-8 Induktive Einflüsse auf C-H-Valenzschwingungen

Gruppierung	symmetrische Valenzschwingungen in cm^{-1}	asymmetrische Valenzschwingungen in cm^{-1}
$ClCH_3$	=2966	=3042
$-O-CH_3$	2955	3020
$-S-CH_3$	2880	2970
$-\underset{|}{\overset{|}{C}}-CH_3$	2870	2960

Tab 5.3.4-9 Induktive Einflüsse auf symmetrische Methyl-Deformationsschwingungen

Gruppierung	Wellenzahl $\delta_s CH_3$ in cm^{-1}
FCH_3	1460
$-OCH_3$	1450
$ClCH_3$	1360
$-\underset{|}{\overset{|}{C}}-CH_3$	1380
$-S-CH_3$	1320
$-P-CH_3$	1300
$-Si-CH_3$	1260
$SnCH_3$	1200
$ZnCH_3$	1170

Für die N—H-Schwingungen könnte man die Wirksamkeit der gleichen Gesetze zeigen. Amid-N-H-Frequenzen liegen z. B. höher als die von Amin-NH, in Übereinstimmung mit der Coplanarität von CNH_2 im ersteren Falle, die eine sp^2-Hybridisierung, also höheren s-Anteil in der NH-Bindung voraussetzt. Es sei bei diesem Element auf eine weitere Gesetzmäßigkeit hingewiesen: *Frequenzerniedrigung mit steigender Bindigkeit.*

$$NH_3 \quad \tilde{\nu}_{as}NH_3 = 3443\ cm^{-1} \quad \tilde{\nu}_s NH_3 = 3336\ cm^{-1}$$

$$NH_4^+ \quad \tilde{\nu}_{as}NH_4 = 3314\ cm^{-1} \quad \tilde{\nu}_s NH_4 = 3233\ cm^{-1}$$

$$R{-}NH_2 \quad \tilde{\nu}_{as}NH_2 = 3390\ cm^{-1} \quad \tilde{\nu}_s NH_2 = 3520\ cm^{-1}$$

$$R{-}NH_3^+ \quad \left.\begin{matrix}\tilde{\nu}_{as}NH_3\\ \tilde{\nu}_s NH_3\end{matrix}\right\} \approx 3000\ cm^{-1}$$

Danach hat man aus den Schwingungsspektren zuerst entnehmen können, daß die Aminosäuren in kristallisierter Form und in etwa neutraler Lösung als Zwitterionen vorliegen: $H_3N^{\oplus}{-}CH_3{-}COO^{\ominus}$.

Die universelle Gültigkeit dieser Gesetze zeigt sich darin, daß auch die eventuell unterschiedlichen Koordinationszahlen von Kationen danach erkannt werden können (Tab. 5.3.4-10).

Die N-H-Schwingungen werden durch Effekte beeinflußt, die bei OH-Gruppierungen stärker auftreten und das Aussehen der Spektren entscheidend beeinflussen: die Assoziation über Wasserstoffbrücken OH...O bzw. NH...O oder allgemein X-H...Y. Die spektroskopischen Effekte der Wasserstoffbrücken sind die Erniedrigung der X-H-Valenzschwingungen wegen der Lockerung der X-H-Bindungen. Gleichzeitig tritt eine Intensivierung der Infrarotabsorption und eine Verbreiterung wegen des mehr polaren Bindungszustands ein. Zwischen freien und assoziierten Alkoholmolekülen beträgt der Wellenzahlunterschied 3630 bis 3300 cm^{-1} $\triangleq$ 330 cm^{-1}.

Das Ausmaß der Wellenzahlerniedrigung ist ein Maß für die Festigkeit der Wasserstoffbrückenbindung, genauer als der röntgenographisch bestimmbare Abstand zwischen den Atomen X und Y.

Tab. 5.3.4-10 Metall-Sauerstoff-Frequenzen in cm^{-1} in Abhängigkeit von der Koordinationszahl (Mittelwerte)

Kation	Koordination	
	tetraedrisch	oktaedrisch
Li^+	530	300
Mg^{2+}	550	370
Al^{3+}	720	470
Fe^{3+}	620	420
Ti^{4+}	720	550

Eine gewisse Frequenzerniedrigung findet sich auch für die Valenzschwingung des Akzeptoratoms Y. Die neue zwischenmolekulare Bindung der Wasserstoffbrücke geht auf Kosten der ursprünglichen Stärke der innermolekularen Bindung.

Durch das Studium der Wasserstoffbrückenabsorptionen in Lösungen von verschiedenen Konzentrationen bei verschiedenen Temperaturen konnten Gleichgewichtskonstanten und Enthalpien für viele Assoziationsvorgänge ermittelt werden. Von Konstitutionsfragen der OH-Verbindungen gelangt man zu deren Thermodynamik.

5.3.4.3.3. *Strukturforschung.* Bei der Besprechung der wesentlichen Gesetzmäßigkeiten für die Identifizierung der Baugruppen eines Moleküls sind schon Punkte berührt worden, die zu Strukturfragen der Gruppen gehören, z. B. die Unterscheidung planaren und pyramidalen Baus bei Stickstoffverbindungen.

Bei kleinen symmetrischen Molekülen kann die räumliche Struktur durch Vergleich zwischen Raman- und Infrarotspektrum oft ermittelt werden. Sind symmetrische dreiatomige Moleküle *linear* gebaut, so gibt es Auswahlregeln, nach denen die symmetrische Valenzschwingung nur im Ramaneffekt, die antisymmetrische Valenzschwingung und die Deformationsschwingung aber nur in Infrarotabsorption zu beobachten sind. Bei *gewinkelten Molekülen* erscheinen alle drei Schwingungen in beiden Effekten. Im Unterschied zu H_2O, SO_2, O_3 sind die Moleküle CO_2, CS_2, N_2O, HCN linear gebaut. N_2O ist aber nicht symmetrisch wie CO_2, sondern hat die Atomanordnung N—N—O. Auch bei dieser Struktur erscheinen drei Grundschwingungen im Infrarotspektrum und im Ramaneffekt, wie auch bei HCN. In diesen Fällen folgt der lineare Bau auch aus der Rotationsstruktur der Infrarotbanden.

Ebenfalls nach symmetriebedingten Auswahlregeln ergibt sich für *Benzen* aus den Spektren, daß alle C—C-Bindungen gleich sein müssen, so daß tatsächlich ein vollkommener Bindungsausgleich eintritt.

Durch die optische Spektroskopie wurde erstmals überzeugend dargestellt, daß Atomgruppen um Einfachbindungen herum leicht drehbar sind und daß die Mehrzahl aller Kettenmoleküle deshalb aus Gemischen von verschiedenen *Rotationsisomeren* besteht. Die Unterschiede zwischen den Spektren der einzelnen Rotationsisomeren sind zwar gering, es läßt sich aber in einfacheren Fällen verfolgen, wie die eine oder andere Form bei Änderung des Aggregatzustands oder unter der Wirkung von Lösungsmitteln sich bevorzugt ausbildet. Es ist davon

auszugehen, daß gewöhnlich schon aus statistischen Gründen die eine Form häufiger ist als die andere. Die Darstellung der Rotationsisomerie für 1,2-Dichlorethan in der Newmanprojektion zeigt, daß es für die gauche-Form zwei Realisierungsmöglichkeiten gibt (Abb. 5.3.4-11). Spektroskopisch sind diese beiden

trans-1,2-Dichlorethan gauche-1,2-Dichlorethan

Abb. 5.3.4-11 Rotationsisomerie für 1,2-Dichlorethan in der Newmanprojektion

optischen Isomere nicht zu unterscheiden. Sie sind auch nicht trennbar, da sie sich leicht durch die wenig behinderte Drehung um die C—C-Bindung ineinander und in die trans-Form umwandeln. Im Gaszustand ist tatsächlich die gauche-Form in der doppelten Konzentration wie die trans-Form nachzuweisen. In der flüssigen Phase ist diese zusätzlich benachteiligt. Sie hat infolge der hohen Symmetrie kein elektrisches Dipolmoment. Bei der gauche-Form ergibt sich eine Dipol-Dipol-Anziehung zwischen den Molekülen, eine Art sehr schwacher Verbindungsbildung mit Energieabgabe. Diese Form wird damit zusätzlich begünstigt. Im festen Zustand liegt aber allein das trans-Isomer vor. Es ist für den Einbau in den Kristall geeigneter. Kristallisation wird überhaupt nur dadurch möglich, daß die Moleküle dieses Stoffs sich in die trans-Form umlagern. Diese Strukturfragen der Konformation wurden zuerst und werden erfolgreich mit Infrarot- und Ramanspektroskopie bearbeitet, heute besonders auch bei der Erforschung der Hochpolymere.

5.3.4.3.4. *Bindungskraftkonstanten.* Aus den Molekülschwingungen lassen sich die *Kraftkonstanten* berechnen und damit *Festigkeiten der chemischen Bindungen* untersuchen. Zwar ist auch aus den intramolekularen Atomabständen, die durch röntgenographische Untersuchungen oder durch Elektronenbeugung ermittelt werden, ein Schluß auf die Bindefestigkeit möglich, weil bei stärkeren Bindungen kürzere Atomabstände auftreten. Die Effekte sind dabei aber nur klein. Außerdem kann man in der Spektroskopie ohne jede Rechnung als grobe Näherung eine charakteristische Frequenz bereits als ein Maß für die Festigkeit der Bindung benutzen.

Einfachbindungen zwischen leichteren Atomen liegen immer in der Nähe von 1000 cm^{-1}, die Schwingungen von Zweifachbindungen bei 1500 cm^{-1} und die der Dreifachbindungen um 2000 cm^{-1}. Schon daraus ergibt sich, daß das Sauerstoffmolekül eine wirkliche Doppelbindung hat, wenn diese auch unter Mitwirkung von Elektronen mit parallelen Spins gebildet wird. (↑ 5.1.2.). Kohlenoxid hat ganz eindeutig eine Dreifachbindung, wie sie in der Elektronenschreibweise |C≡O| zum Ausdruck gebracht wird.

Ohne Ausrechnung der Kraftkonstanten ist es auch anhand der Frequenzen möglich, das Vorliegen von Doppelbindungen in den Säuren, z. B. der Elemente der 3. Periode, nachzuweisen und den mesomeren Bindungsausgleich in den Anionen zu verfolgen. Für Schwefelsäure und Phosphorsäure sind demnach die älteren Formeln, bei denen die *Valenzstriche* einfach als Symbole der Wertigkeit aufgefaßt werden, auch richtig, wenn man den Strichen die gegenständliche Bedeutung von Elektronenpaaren beilegt. Die Schwingungsspektren zeigen, daß sich im Sulfatanion $1^1/_2$fach-Bindungen, im Hydrogensulfatanion aber $1^1/_2$-Bindungen neben einer S—O-Einfachbindung befinden (Abb. 5.3.4-12),

Sulfatanion

$\nu_{as}\, SO_4^{2-} = 1104\, cm^{-1}$

$\nu_s\, SO_4^{2-} = 981\, cm^{-1}$

Hydrogensulfatanion

$\nu_{as}\, SO_3^{-} = 1200\, cm^{-1}$

$\nu_s\, SO_3^{-} = 1051\, cm^{-1}$

$\nu\, S{-}OH = 887\, cm^{-1}$

Abb. 5.3.4-12 Bindungen und Valenzschwingungswellenzahlen im Sulfat- und Hydrogensulfatanion; der Kreis oder Kreisbogen wird zur bildlichen Darstellung des mesomeren Bindungsausgleichs um das Zentralatom herum verwendet

Stickstoff kann nicht 12 Elektronen in seiner Valenzschale unterbringen. Aus den Schwingungswellenzahlen des Nitratanions oder der Nitroverbindungen folgt, daß nur eine Doppelbindung beteiligt sein kann, wie bei entsprechenden Carbonat- oder Carbonsäureanionen:

$\tilde{\nu}_{as}CO_3^{2-} = 1415\ cm^{-1}$

$\tilde{\nu}_{s}CO_3^{2-} = 1063\ cm^{-1}$

$\tilde{\nu}_{as}NO_3^{-} = 1390\ cm^{-1}$

$\tilde{\nu}_{s}NO_3^{-} = 1050\ cm^{-1}$

$\tilde{\nu}_{as}CO_2^- = 1552\ cm^{-1}$
$\tilde{\nu}_{s}CO_2^- = 1416\ cm^{-1}$

$\tilde{\nu}_{as}NO_2 = 1476\ cm^{-1}$
$\tilde{\nu}_{s}NO_2 = 1345\ cm^{-1}$.

Bei exakter Berechnung der Kraftkonstanten geben die Schwingungsspektren auch noch Auskunft über kleinere Veränderungen in den Festigkeiten der chemischen Bindungen, die sich in den Valenzstrichformeln nicht darstellen lassen. Die Bindekraft der Atome in einem Molekül wird durch ihre weiteren Nachbarn beeinflußt. So berechnet man z. B. die folgenden Kraftkonstanten für die verschiedenen C—Cl-Bindungen in den Molekülen. Es ergibt sich für CCl_4 f = 312 Nm^{-1}; für $CHCl_3$ f = 300 Nm^{-1}; für $Cl_2HC—CHCl_2$ f = 321 Nm^{-1}; für CH_3Cl f = 342 Nm^{-1}.

5.3.4.4. Kopplung der Spektrometer mit elektronischen Rechenmaschinen. Es gibt eine Anzahl von spektroskopischen Untersuchungen, die erhebliche rechnerische Auswertungen erfordern. Die Bestimmung der Konzentrationen in einem *Mehrkomponentensystem* mit Hilfe vorher durch Eichmessungen bestimmter Extinktionskoeffizienten z. B. erfordert für jede Probe die Auflösung eines Gleichungssystems der Art

$$\begin{aligned} E_1/l &= \varepsilon_{1A}c_A + \varepsilon_{1B}c_B + \varepsilon_{1C}c_C \\ E_2/l &= \varepsilon_{2A}c_A + \varepsilon_{2B}c_B + \varepsilon_{2C}c_C \\ E_3/l &= \varepsilon_{3A}c_A + \varepsilon_{3B}c_B + \varepsilon_{3C}c_C \end{aligned} \quad (6)$$

mit drei bei Wellenlängen λ_1, λ_2 und λ_3 mit Schichtdicke l gemessenen Extinktionen $E_{1,2,3}$, den Extinktionskoeffizienten $\varepsilon_{1,2,3}$ bei den Wellenlängen λ_1, λ_2 und λ_3, jeweils für die reinen Stoffe A, B und C, und deren Konzentrationen $c_{A,B,C}$ als Unbekannten bei Vorliegen eines Dreistoffgemischs; wie es bei noch mehr Komponenten aussieht, ist vorstellbar. Der Ansatz des Gleichungssystems erfordert als erstes die Umrechnung der primär meist gemessenen Durchlässigkeiten in Extinktionswerte. Dann ist nach den Unbekannten $c_{A,B,C}$ aufzulösen.

Erheblichen Rechenaufwand erfordert weiter die statistische Bearbeitung von Analysenergebnissen. Sie geht aus von dem *Mittelwert* eines Ergebnisses x_i bei m Bestimmungen:

$$\bar{x} = (x_1 + x_2 + \ldots + x_m)/m \quad (7)$$

und erfordert weiter die Berechnung der *Standardabweichung* s:

$$s = \sum_{i=1}^{m} (\bar{x} - x_i)^2/m, \quad (8)$$

eine Auswertung, die bei analytischen Arbeiten wegen zu hohen Rechenaufwands gewöhnlich unterbleibt. So hat sich von den analytischen Aufgaben her die weitere Bearbeitung von spektrometrischen Messungen durch Rechenmaschinen ergeben.

Ausgehend von den Spektrometern üblicher Konstruktion läßt sich eine off-line-Kopplung mit einem Rechner bewirken, wenn eine Datenausgabe auf Lochstreifen geschaffen wird. Die Lochstreifen können dann einem beliebigen Rechner (der den Code der Streifen verarbeiten kann) zugeführt werden. Es ist dabei erforderlich, das in *analoger Form* für die Spektrenregistrierung vorliegende Signal durch einen Analogdigitalumsetzer zu digitalisieren, und des weiteren eine digitalisierte Form der Wellenlängen bzw. Wellenzahlanzeige zu schaffen. Erfahrungen mit numerisch gesteuerten Werkzeugmaschinen legen eine engere Kopplung zwischen Spektrometer und Rechner nahe, bei der auch die Steuerung von Spektrometerfunktionen vom Rechner aus erfolgt. Wenn Messungen für das Gleichungssystem 5.3.(6) wiederholt an den Meßstellen λ_1, λ_2 und λ_3 erfolgen müssen, kann der Rechner sehr gut die Einstellung des Spektrometers durch entsprechende Befehle steuern. Die Bedienschalter am Spektrometer müssen gleichzeitig als Relais ausgebildet sein, die auf Befehle vom Rechner ansprechen. So sind heute Gerätesysteme gebräuchlich geworden, bei denen mehrere Spektrometer mit einem Kleinsteuerrechner verbunden sind. Die spektroskopischen Methoden sind besonders gut für eine derartige Laboratoriumsautomatisierung geeignet. Auch in der qualitativen Analyse, d. h. bei der Identifizierung unbekannter Substanzen nach ihren Spektren, werden mehr und mehr Rechenmaschinen eingesetzt. Es werden zwei Wege beschritten. Einmal ist es möglich, nach einem Rechenprogramm all die Schlüsse vollziehen zu lassen, die ein Spektroskopiker nach empirischen Regeln und theoretischen Gesetzen bei der Interpretation eines Substanzspektrums systematisch benutzen würde, um die Art und vielleicht die Identität der Substanz zu erkennen. Andererseits kann man aber auch einen Computer eine riesige Menge bekannter Spektren, die in vereinfachter Form im Speicher codiert eingelesen sind, mit dem Spektrum einer unbekannten Verbindung sehr schnell vergleichen lassen. Wenn das Spektrum der unbekannten Substanz nicht gespeichert ist, werden – wenn das System gut ist – ähnliche Substanzen mit graduell geringerer Übereinstimmung ausgedruckt, die zur endgültigen Identifizierung hinführen.

Mit der Rechnerkopplung konnte die Präzision der spektroskopischen Messungen verbessert und mehr Information aus den Spektren abgeleitet werden. Man kann Spektren oder einzelne Banden mehrfach messen und Mittelwertbildungen für Spektralbereiche und ganze Spek-

tren vornehmen; denn an sich ist jedes Spektrum durch geringfügige statistische Schwankungen im optischen oder elektronischen Teil der Apparatur etwas verrauscht. Ist dieses *Rauschen* statistisch, so verschwindet es bei Mittelwertsbildung. Effektiver noch sind mathematische *Glättungsoperationen*, bei denen etwa vorausgesetzt wird, daß jeder Meßpunkt nahe bei dem Mittelwert seiner Nachbarpunkte liegen sollte. Für Extrema gilt dies nicht, und so werden in Wirklichkeit kompliziertere Funktionen angewandt. Jedenfalls wird jeder einzelne Punkt durch Berechnung während der Spektrenregistrierung oder nachträglich nachgeprüft und korrigiert. Erst mit derartig verbesserten Spektren ist es lohnend, weitere Operationen vorzunehmen: Maßstabsveränderungen, um evtl. unsichtbare Absorptionen bei starker Vergrößerung darzustellen, Spektren von Lösungsmitteln oder Hauptkomponenten zu subtrahieren, um Spektren von Nebenbestandteilen zu isolieren, u. ä. Die im Falle der Fouriertransformationsspektroskopie notwendigen Umformungen gehören auch zu diesen Rechneranwendungen.
Eine direkte Einführung von Rechnerfunktionen in ein Spektrometer erlauben die *Mikroprozessoren*. Diese Elemente der Mikroelektronik können primäre Umformungen der Daten bei der Messung vornehmen und bei der Steuerung der Spektrometer Getriebe oder Kurvenscheiben ersetzen, deren mechanische Herstellung sehr arbeitsaufwendig ist. Weitergehende Berechnungen nach von Fall zu Fall wechselnden Programmen müssen äußeren angekoppelten Rechnern überlassen bleiben.

5.3.5. Mikrowellenspektroskopie

5.3.5.1. Mikrowellenerzeugung und -nachweis. Die Frequenzen der Molekülrotation entsprechen denen der kürzesten heute mit den Mitteln der Hochfrequenztechnik erzeugbaren Radiowellen. Für Radarzwecke wurde etwa ab 1940 die *Dezimeterwellentechnik* entwickelt und später auf die Erzeugung der noch kürzeren *Mikrowellen* ausgedehnt. Da deren Wellenlängen in der Größenordnung von Zentimetern bis Zehntelmillimetern den Abmessungen der verwendeten Apparate entsprechen, können Mikrowellen durch passend geformte Rohre geleitet werden, während sich längere Radiowellen nach allen Seiten in den Raum hinaus ausbreiten und sich nur durch Reflektoren richten lassen. Dieses Zusammenhalten in Rohren ist wegen der erzeugten kleinen Leistung unentbehrlich. Die *Hohlleiter* haben kreisförmigen oder rechteckigen Querschnitt. Ihr Durchmesser begrenzt den Energietransport; wird die Weite der halben Wellenlänge gleich, so bilden sich stehende Wellen aus, und bei weiterer Verengung wird die Schwingungsenergie durch Dämpfung vernichtet. Durch die Metallwände soll keine elektrische Ladung abfließen, sie dienen nicht der Stromleitung, sondern nur der Begrenzung des Wellenfelds. In einem *Hohlraumresonator* wird an bestimmten Stellen durch passende Dimensionierung die Energie in einer stehenden Welle konzentriert. Er ist gleichzeitig ein wichtiges Bauelement der Spezialsenderöhren für Mikrowellen. In dem *Reflexionsklystron* (Abb. 5.3.5-1) wird ein Elektronenstrahl

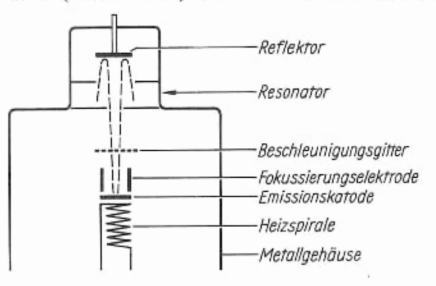


Abb. 5.3.5-1 Reflexionsklystron

gegen eine positiv aufgeladene Reflektorelektrode gerichtet und durch das starke Gegenfeld vollständig umgelenkt. Diese Elektronen treten dann durch schmale Schlitze in einen Hohlraumresonator ein. Bei passender Geschwindigkeit der Ladungszufuhr gerät die Ladungswolke in Schwingung. Durch die Abstoßung der nachfolgenden Elektronen überträgt sich der Rhythmus auf den Strahl, so daß die weitere Zufuhr in Schüben erfolgt. Die Schwingung wird immer weiter verstärkt bzw. aufrechterhalten, falls Energie aus dem Resonator abfließt. Das Reflexionsklystron geht auf die Bremsfeldröhre von Barkhausen und Kurz zurück. Auch andere Methoden der Mikrowellenerzeugung sind für spektroskopische Zwecke verwendbar.
Durch die Resonatordimensionen ist die Röhre auf eine konstante Wellenlänge abgestimmt. Sie läßt sich durch Veränderung der Reflexionsspannung verstimmen und arbeitet dann unter ungünstiger Leistungsabgabe, die ohnehin nur einige 0,01 W beträgt. Durch Veränderung des Resonatorvolumens mit Hilfe einer beweglichen Membran oder durch andere Vorrichtungen (im Hochvakuum) kann man die Klystronfrequenz um $\pm 20\%$ verstellen.
Für eine Absorptionsuntersuchung muß deshalb ein ganzer Satz von Senderöhren zur Verfügung stehen, damit ein Stoff mit einer kontinuierlichen Folge von Frequenzen der Reihe nach durchstrahlt werden kann. Andererseits entfällt für diese Art von Spektroskopie das dispergierende Element, ein Gitter bzw. ein Prisma.

Als *Empfänger* dienen *Kristalldetektoren.* Statt der Bleisulfidkristalle, die in den Anfängen der Hochfrequenztechnik als Radioempfänger dienten, benutzt man meist Siliciumkristalle, auf denen die Spitze eines Wolfram-Haarkristalls ruht. Die jeweils eingestrahlte Frequenz kann auf 10^{-5}%, d. h. bei 10^5 MHz auf ±0,01 MHz genau gemessen werden, wenn man die Schwebungen mit der Frequenz eines Eichsenders beobachtet, der von einem piezoelektrischen Kristall gesteuert wird. Infolge dieser hohen Meßgenauigkeit ist die Auflösung in der Hochfrequenzspektroskopie besser als bei jeder anderen spektroskopischen Methode. In einem Bereich von 0,1 bis 10 cm^{-1}, d. h. von 300000 bis 3000 MHz, könnten rund 50 Mill. Spektrallinien voneinander unterschieden werden. In der Mikrowellenspektroskopie ist wie in der Hochfrequenzspektroskopie insgesamt die Angabe von Frequenzen üblich.

5.3.5.2. Untersuchbare Stoffe. Es treten nur Stoffe mit einem *elektrischen Dipolmoment* mit den Mikrowellen – ebenso wie mit der infraroten Strahlung und jeder elektromagnetischen Strahlung überhaupt – durch Absorption in Wechselwirkung. Die gleichatomigen zweiatomigen Moleküle scheiden deshalb auch hier aus. Die Nachweismöglichkeiten sind im Laufe der Zeit immer mehr gesteigert worden. Man hat bei Deuteroacetylen H—C≡C—D, einem eigentlich unpolaren Molekül, Linien beobachten können ($J = 0 \rightarrow 1$ bei 59450 MHz), weil nur infolge geringer Unterschiede in der C—H- und C—D-Bindungslänge ein minimales Dipolmoment entsteht. (Der Unterschied in den C—H- und C—D-Bindungslängen rührt von den unterschiedlichen Nullpunktsamplituden von H und D her, vgl. 5.3.1.2.).

Paramagnetische Moleküle können wegen ihres magnetischen Moments in Wechselwirkung mit den elektromagnetischen Wellen treten; z. B. ist ein Mikrowellenspektrum von Sauerstoff beobachtet worden. Bei Stoffen, die Wasserstoff und nicht wenigstens zwei schwere Atome enthalten, liegt wegen der dann großen Rotationskonstanten schon der niedrigste Übergang (≈10 cm^{-1}) an der Grenze des Mikrowellenbereichs. Einige Substanzen mit den einfachsten Spektren können deshalb nicht erfaßt werden. Untersuchungen über *Rotationslinien* sind nur an gasförmigen Stoffen möglich. Um die auch bei Gasen noch bestehende Druckverbreiterung der Linien zurückzudrängen, die durch die Stöße der Moleküle untereinander und die dabei erfolgende Störung der Rotationsenergien entsteht, wird bei möglichst geringem Gasdruck von etwa 10^{-5} bis 10^{-3} Pa gearbeitet. Je nach dem Absorptionsvermögen sind dann unter Umständen Rohre von 10 m Länge erforderlich. Für Stoffe, die unter Normalbedingungen flüssig oder fest sind, hat man Heizzellen für Temperaturen bis 1300 K gebaut; wenn keine Zersetzung eintritt, kann dann meist ein Dampfdruck von 10^{-5} Pa erzielt werden.

Die Mikrowellenabsorption ist der Energie nach ein schwacher Effekt; für intensivere Absorptionen erreicht man mit 3 m Zellenlänge an der Stelle einer Absorptionsfrequenz eine Schwächung der eingestrahlten Energie um rund 15%. Dennoch können mit Kunstgriffen millionenfach schwächere Linien registriert werden.

5.3.5.3. Hyperfeinstruktur. Bei der großen erzielbaren Auflösung kann außer den Übergängen zwischen den Rotationsniveaus eine Hyperfeinstruktur der Rotationslinien beobachtet werden. Diese beruht auf einer Wechselwirkung zwischen der Molekülrotation mit den Spinmomenten der Atomkerne, sofern diese Kernspinmomente haben (↑ 4.4.2.). Die Rotation des Moleküls erzeugt ein äußerst schwaches Magnetfeld (einige 10^{-2} Am^{-1}). Damit koppeln die ebenfalls kleinen Kernspinmomente nach Quantengesetzen, also unter Multiplettbildung. Diese Hyperfeinstruktur, die vom magnetischen Kernmoment herrührt, wird gewöhnlich nicht aufgelöst und ist eher eine störende Komplikation als ein Effekt, der zusätzliche Informationen liefert.

Um einige Größenordnungen stärker kann eine Multiplettaufspaltung der Rotationslinien ausfallen, die von den elektrischen Quadrupolmomenten gewisser Atomkerne (↑ 4.4.6.) herrührt. Dabei wirkt ein zweiter Faktor mit: Der Gradient der elektrischen Feldstärke am Ort des Quadrupolkerns bestimmt die Größe der Aufspaltung $E = eQ/hc \; d^2V/dz^2 \, f(I, J)$, genauso wie in der Kernquadrupolresonanzspektroskopie (↑ 4.4.6.). Der Faktor $f(I, J)$ enthält in komplizierter Weise die Kernspin- und Rotationsquantenzahl. Das Multiplett ergibt sich daraus, daß es für das Gesamtmoment, dessen Größe durch die Quantenzahl F dargestellt wird, die Zustände $F = J + I$, $F = J + I - 1, \ldots,$ $F = J - I$ geben kann. Für $J = 0$ gibt es keine Aufspaltung, da kein Drehmoment vorhanden ist. An Auswahlregeln gilt neben $\Delta J = \pm 1$: $\Delta F = 0, \pm 1$. So entsteht für den Übergang $J = 0 \rightarrow 1$ in HCN mit dem Kernspinmoment I(N-14) = 1 ein Triplett. Wenn man in dem betrachteten Molekül H durch D, also durch einen weiteren Quadrupolkern ersetzt, wird keine weitere Komplizierung der Multipletts beobachtet. Das Quadrupolmoment von D bleibt wirkungslos, offenbar besteht kein elektrischer Feldgradient am Ort des Wasserstoffatoms. Es befindet sich in einer σ-Bindung zum Kohlenstoff und damit in einer recht gut kugelsymmetrischen Elektronenumgebung.

Die vom Kernspin jederzeit ausgehende magne-

tische Kopplung an die Molekülrotation ist für HCN durch ein Experiment besonders hoher Auflösung, durch Einstrahlung mittels eines Masers, d. h. eines Mikrowellenlasers, sichtbar geworden. Die kleinste Aufspaltung im magnetischen Triplett, in das jede Linie vom Quadrupol-Triplett aufgelöst wird, betrug für eine N-14-Linie bei 72414 MHz 23 kHz. Das bedeutet eine Meßgenauigkeit von $10^8 : 1$, entsprechend der Aufgabe, 1 km auf 0,01 mm genau auszumessen. Es ist üblich, statt von magnetischer Dipolkopplung und elektrischer Quadrupolkopplung kurz von magnetischer und von Quadrupolkopplung zu sprechen.

Die Quadrupolhyperfeinstruktur ist ein kompliziertes Phänomen gegenüber der einfachen Kernquadrupolresonanz (↑ 4.4.6.), weil die Rotationszustände der Moleküle mit eingehen. Man untersuchte bisher meist nur Fälle, in denen nur ein Quadrupolkern mitwirkt. Die elektronischen Rechenmaschinen erlauben aber die Vorausberechnung beliebig komplizierter Spektren.

5.3.5.4. Molekularer Starkeffekt. Durch eine zweite Quantenzahl m sind die $(2J + 1)$ verschiedenen Orientierungsmöglichkeiten der Rotationsachse im Raum abzählbar. Diese verschiedenen Orientierungen gehören im feldfreien Raum zur gleichen Energie; für die Intensität der Rotationslinien muß diese $(2J + 1)$-fache *Entartung* berücksichtigt werden. In einem elektrischen Feld sind aber die verschiedenen Stellungen der Molekülachsen nicht mehr energiegleich, die Entartung wird aufgehoben. In Analogie zu den Atomspektren wird ein molekularer Starkeffekt beobachtet. Er wird erzeugt, indem durch die Gaszelle zentral ein gegen die Wände isolierter Draht gezogen wird, an den eine elektrische Hochspannung gelegt ist, oder besser durch eine isolierte ebene Zwischenwand im rechteckigen Hohlleiter. Neben der Aufspaltung tritt eine Verschiebung – auch für den nicht aufspaltenden rotationslosen Grundzustand – auf.

Der Starkeffekt ist ein Hilfsmittel zur Erhöhung der Meßempfindlichkeit und für die richtige Interpretation der Mikrowellenspektren. Zum ersteren Zweck wird das Starkfeld durch Wechselstrom aufgebaut. Die Moleküle unterliegen dem Effekt periodisch. Eine bestimmte Linie taucht bei einer bestimmten Mikrowellenfrequenz auf und verschwindet wieder im raschen Wechsel. Damit empfängt der Detektor ein Wechselstromsignal, welches leichter und unter Anwendung von Bandfiltern verstärkt werden kann. Für die Zuordnung zu den einzelnen Rotationsquantenzahlen ist es von Wert, daß ein Starkmultiplett immer aus $2J + 1$ Linien besteht. Die Aufspaltung nimmt mit wachsendem J ab.

5.3.5.5. Molekularer Zeemaneffekt. In Analogie zum Verhalten der Atome im Magnetfeld (↑ 4.2.4.6.) gibt es auch einen molekularen Zeemaneffekt, der für die paramagnetischen Moleküle O_2, NO, NO_2 und ClO_2 zu ähnlichen Multipletts wie der Starkeffekt führt. Schon das Magnetfeld der Erde reicht aus, um meßbare Linienaufspaltungen herbeizuführen. Andererseits hat jedes andere Molekül einen minimalen Restparamagnetismus. Durch Anwendung sehr starker Felder gelingt es, auch bei den als diamagnetisch bezeichneten Molekülen geringe Zeemanaufspaltungen hervorzubringen.

5.3.5.6. Ergebnisse der Mikrowellenspektroskopie. Durch die genaue Vermessung der Rotationslinien können Trägheitsmomente und damit *Atomabstände* und *Bindungswinkel* mit größerer Präzision als nach irgend einer anderen Methode ermittelt werden. Je nach der Kreiselsymmetrie des Moleküls sind allerdings von vornherein entweder primitive oder höchst komplizierte Spektren zu erwarten. Bei linearen Molekülen wirkt nur eine Rotationskonstante mit, die im Durchschnitt die Größenordnung 5000 MHz hat, so daß man, von etwaiger Hyperfeinstruktur abgesehen, alle 10000 MHz eine Linie erwartet, d. h. rund 30 Linien in konstanten Abständen im Bereich von 0,1 bis $10\ cm^{-1}$. Erheblich linienreicher, aber von gleicher Systematik wird das Spektrum eines symmetrischen Kreisels. Beim asymmetrischen Kreisel aber gibt es keine Auswahlregeln, so daß in einem Bereich von 10000 MHz anstatt einer Linie eines linearen Moleküls oder eines symmetrischen Kreisels bei weniger symmetrischen Molekülen viele hundert Linien liegen.

Aus einem Trägheitsmoment allein kann man nur für zweiatomige Moleküle direkt den *Atomabstand* berechnen. Man erhält aber zusätzliche Bestimmungsgleichungen, wenn man das gleiche Molekül mit verschiedenen Isotopen untersuchen kann. Trotz gleicher Bindungskräfte und deshalb auch gleicher Atomabstände erhält man infolge der verschiedenen Atommassen ganz verschiedene Rotationslinien. In langen Absorptionsrohren genügt oft schon die natürliche Isotopenverteilung, z. B. bei Schwefel, Kohlenstoff oder Sauerstoffverbindungen, um auch die schwach vertretenen Kombinationen mit Nebenisotopen zu beobachten.

Für einige einfache Moleküle war die Identifizierung in geringer Konzentration oder der Nachweis des linearen Baus von Bedeutung. $H{-}C{\equiv}P$ als Analoges von HCN wurde durch Mikrowellenspektroskopie nachgewiesen. $H{-}C{\equiv}N{-}\overline{\underline{O}}|$ und $|C{\equiv}N{-}\overline{\underline{O}}{-}H$, die Knallsäure und Isoknallsäure, sind lineare Moleküle. Die Isocyansäure ist es nicht infolge einer mesomeren Bindungsverteilung $|\overline{\underline{O}}{-}C{\equiv}N{-}H \leftrightarrow \langle O{=}C{=}\overline{N}{-}H$, so daß wieder die Regel

erkennbar ist, nach der nichtbindende Elektronenpaare wie hier am Stickstoff gewinkelte Struktur bedingen. Im scheinbaren Gegensatz zu dieser Regel hat die Mikrowellenspektroskopie für KOH, RbOH und CsOH lineare Struktur ergeben. Auch aus den Metall-Sauerstoff-Bindungslängen folgt aber, daß hier im wesentlichen Ionenbindungen vorliegen.

Mit Hilfe des *molekularen Starkeffekts* (s. a. 5.3.5.4) sind die *Dipolmomente* der Moleküle bestimmbar. Schon für den rotationslosen Grundzustand tritt eine Verschiebung ein, die dem

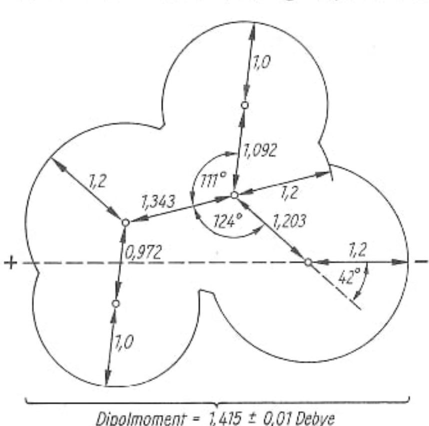


Abb. 5.3.5-2 Struktur und Dipolmoment des Ameisensäuremoleküls HCOOH nach KIM, KELLER und GWINN

Quadrat der Feldstärke und des Dipolmoments proportional ist. Der Effekt ist allerdings klein; die Linie $J = 0 \rightarrow J = 1$ ist um 0,4 MHz gegen den feldfreien Fall verschoben, wenn $E = 300\ \mathrm{Vcm^{-1}}$, $B = 0{,}5\ \mathrm{cm^{-1}}$ und $\mu = 1/3 \times 10^{-29}$ cm betragen. Als Beispiel für die genaue Vermessung eines Moleküls mittels Mikrowellen kann die Untersuchung der Ameisensäure dienen (Abb. 5.3.5-2).

Hyperfeinstrukturmessungen können zu Aussagen über die Elektronenverteilung im Molekül benutzt werden. Sie zeigen z. B. für Alkalihalogenide in der Dampfphase innerhalb der Meßgenauigkeit verschwindende Feldinhomogenitäten an den Kernorten. Das bedeutet, daß die abgeschlossenen Elektronenschalen des Ions Cl^-, dessen Kern ein Quadrupolmoment hat, durch die Bindung nicht wesentlich verzerrt werden. Dagegen zeigen Messungen an Thalliumchlorid, daß die Elektronenhülle des Cl^- dort schon so weit verzerrt wird, daß eine deutliche Feldinhomogenität entsteht.

Übergänge besonderer Art, die im Mikrowellenspektrum erscheinen, sind die *Inversionsübergänge pyramidenförmiger Moleküle*. Es handelt sich dabei nicht um den Energiebetrag, der zum „Umklappen“ des Moleküls selbst notwendig und ein Mehrfaches der Schwingungsquanten ist. Bis z. B. bei NH_3 ein Hindurchschwingen des N-Atoms durch die Ebene der drei Wasserstoffatome eintritt, muß die NH_3-Biegeschwingung dreifach angeregt sein. Es gibt aber in der Wellenmechanik bereits eine Einwirkung der Niveaus der umgeklappten Form auf die der an sich identischen Ausgangsform. Die bloße Möglichkeit des Übergangs bei höheren Energien bringt für die tiefer liegenden Niveaus eine *Aufspaltung* mit sich. Sie beträgt für NH_3 im Grundzustand bereits $0{,}64\ \mathrm{cm^{-1}}$, obgleich noch keine Schwingung angeregt ist. Der Übergang zwischen den um diesen Betrag aufgespaltenen Niveaus liegt im Mikrowellenbereich bei etwa 30 MHz. Weil das Molekül sich gleichzeitig in verschiedenen Rotationszuständen befinden kann, entsteht ein ganzes Spektrum. Diese Absorptionen wurden bereits 1934 beobachtet. Eine ähnliche Aufspaltung gibt es auch für Moleküle, die in verschiedenen Rotationsisomeren vorliegen können.

5.4. Makromoleküle

5.4.1. Eigenschaften makromolekularer Stoffe

Viele Naturstoffe sowie Plastwerkstoffe und Fasern enthalten Makromoleküle. Diesen verdanken sie ihre besonderen Eigenschaften. Die Kunststoffe haben im Vergleich zu den meisten metallischen Werkstoffen eine geringe Dichte und sind gute thermische und elektrische Isolatoren. Von Säuren und Laugen werden sie kaum angegriffen. Einige Plastwerkstoffe sind durchsichtig wie Glas, aber praktisch unzerbrechlich. Natürlich haben sie nicht nur günstige Eigenschaften, aber ihre typischen Vorteile, zu denen oft auch die Kautschukelastizität zu zählen ist, rechtfertigen es, ihre Verfügbarkeit mit unter die Grundvoraussetzungen unserer gegenwärtigen Technik und Zivilisation zu rechnen.

5.4.2. Aufbau der Makromoleküle

Makromoleküle sind Riesenmoleküle mit Molmassen über $10^4\ \mathrm{gmol^{-1}}$. Die höchsten Werte, die bisher gemessen wurden, liegen bei einigen $10^7\ \mathrm{gmol^{-1}}$. Makromoleküle stellen Gebilde aus verhältnismäßig einfachen Grundbausteinen dar, die durch normale chemische Bindungen miteinander verknüpft sind. Beispielsweise besteht Polyvinylchlorid (PVC) aus C_2H_3Cl-Bausteinen, die über Kohlenstoffatome miteinander verbunden sind. Der Aufbau aus sich wiederholenden Struktureinheiten ermöglicht

es, die riesigen Moleküle verhältnismäßig einfach zu beschreiben. Es ist die Struktur der Grundbausteine anzugeben und die Anzahl der Grundbausteine, die jeweils ein Makromolekül bilden. Letztere Größe nennt man *Polymerisationsgrad P*. Er verbindet die Masse eines Grundbausteins M_0 mit der Molmasse M des Polymers: $M = PM_0$.

5.4.2.1. Makromoleküle mit Kohlenstoffgerüst. Das Bauprinzip der Makromoleküle kann mit allen mehrwertigen Elementen verwirklicht werden. Die makromolekulare Chemie ist daher nicht einfach ein Spezialgebiet der organischen Chemie; da jedoch Kohlenstoffatome in ganz besonderem Maße dazu neigen, sich miteinander auch in großer Zahl zu verbinden, ist dieses Element in den meisten Makromolekülen vertreten.

Zahlreiche Polymere von großer technischer Bedeutung haben eine reine Kohlenstoffkette. Die Polyolefine (Polyethylen, Polypropylen, Polyisobutylen u. a.), Polystyren, Polybutadien und Polyisopren bestehen überhaupt nur aus Kohlenstoff und Wasserstoff. Bei den Vinylpolymeren findet sich daneben noch Sauerstoff, Stickstoff, Halogen oder dergleichen in Seitengruppen.

Ionomere sind Ethylen-Copolymerisate, die seitständige, mit Kationen wie Mg^{++} oder Zn^{++} umgesetzte Carboxylgruppen aufweisen. Die Salzbildung bewirkt Vernetzung. In der Wärme wird diese so weit aufgehoben, daß die Ionomeren thermoplastisch verarbeitet werden können. Bei Zimmertemperatur bildet sie sich wieder aus und verleiht dem Material mechanische Festigkeit, Steife und Gummielastizität.

5.4.2.2. Polymere mit Kohlenstoff und Heteroelementen im Gerüst. *Sauerstoff* tritt neben Kohlenstoff in den Ketten von Polyestern, Polyethern und Polyacetalen auf, im Polyphenylenoxid, den Epoxidharzen sowie in Polysacchariden. *Stickstoff* findet sich im Gerüst der Polypeptide, der Polyamide und -imide, der Polyharnstoffe sowie der Aminoharze. *Stickstoff und Sauerstoff* kommen neben Kohlenstoff im Gerüst der Polyurethane vor. *Schwefel* trifft man neben Kohlenstoff in den Polysulfidelastomeren und den mit Schwefel vulkanisierten Kautschuktypen an.

5.4.2.2.1. *Chelatpolymere*. Chelate oder „Krebsscheren"-Verbindungen sind Komplexe, deren Liganden mit dem Zentralatom sowohl durch eine polare als auch durch eine koordinative Bindung verknüpft sind. Stabile Chelatkomplexe werden von den β-Diketonen gebildet, z. B. von Acetylaceton und seinen Derivaten. Makromoleküle können mit Liganden aufgebaut werden, die den Rest des Acetylacetons zweimal enthalten und so nach zwei Seiten Chelate bilden können. Als Zentralatome kommen Zink, Cobalt, Kupfer, Cadmium, Nickel, Magnesium und Beryllium in Betracht. Die Polymeren besitzen interessante physikalische Eigenschaften, z. B. hohe thermische Beständigkeit (Abb. 5.4.2-1).

Abb. 5.4.2-1 Beryllium-Chelatpolymer (Kettenausschnitt)

5.4.2.3. Elementorganische Polymere. Im Gerüst elementorganischer Makromoleküle folgen elektropositive Elemente wie Bor, Aluminium, Silicium, Titan, Germanium, Phosphor oder Arsen und elektronegative Elemente wie Sauerstoff, Schwefel oder Stickstoff abwechselnd aufeinander. Kohlenstoff ist in der Regel nur in den Seitengruppen enthalten.

Die Bindungsenergien liegen bis zu 50% höher als die für die C—C-Bindung, aber die danach zu erwartende Temperaturbeständigkeit kann wegen Empfindlichkeit gegen chemische Einflüsse nur in bestimmten Fällen genutzt werden, z. B. beim Einsatz im Weltraum, bei dreidimensional vernetzten Polymerstrukturen bzw. bei Abschirmung der anfälligen Bindungen durch Schutzgruppen. Zu den elementorganischen Polymeren gehören die Silicone, die großtechnisch erzeugt werden und wegen ihrer hervorragenden Eigenschaften hohe Bedeutung haben (Abb. 5.4.2-2).

Abb. 5.4.2-2 Kettenausschnitt von Dimethylsilicon

5.4.2.4. Anorganische Polymere. Einige Elemente wie Schwefel, Phosphor, Selen und Tellur bilden polymere Modifikationen. Wird z. B. eine Schwefelschmelze erwärmt, steigt bei 432 K die Viskosität plötzlich um zwei Zehnerpotenzen an, weil S_8-Ringe zu langen Ketten polymerisieren. Durch Abschrecken kann plastischer Schwefel erhalten werden, der bei starker Dehnung faserförmig erstarrt. In

ähnlicher Weise stellt die Umwandlung von weißem Phosphor in die rote oder schwarze Modifikation eine Polymerisation von P_4-Monomeren dar.
Technische Bedeutung haben die Polyphosphate mit —P—O-Ketten und die Silicatgläser, die wie die Silicone —Si—O-Ketten enthalten. Die übrigen Valenzen am Silicium sind in diesem Fall durch Ionenpaare besetzt, z. B. durch die Gruppierungen $—O^-Na^+$ beim Natriumsilicat. Auch die Kettenenden werden so gebildet. Der Polymerisationsgrad liegt daher um so niedriger, je höher der Na_2O-Gehalt des Glases ist (s. a. 6.4.6.).

5.4.3. Größe der Makromoleküle

Bei Polymeren ist die Kenntnis der Molmasse bzw. des Polymerisationsgrads von fundamentaler Bedeutung (↑ 5.4.2.).
Die Molmassebestimmungen sind hier jedoch weitaus problematischer als bei niedermolekularen Stoffen.
5.4.3.1. Molekülbegriff. In der niedermolekularen Chemie sind alle Moleküle einer Verbindung gleich groß und gleich gebaut. Bei Polymeren ist dagegen die Zahl der Grundbausteine pro Makromolekül in der Regel statistischen Gesetzen unterworfen. Außerdem treten Abweichungen von der Idealstruktur auf, wie Verzweigungen oder Doppelbindungen, die ebenfalls statistisch verteilt sind. Jedoch gilt auch für Makromoleküle, daß alle Atome im Molekül durch normale Hauptvalenzbindungen miteinander verknüpft sind.

	Monomer	Oligomere
molare Masse $gmol^{-1}$	~100	100 < M ⩽ 10000
Beispiel	$CH_2{=}CH$ (Phenyl)	$CH_2{=}C{-}CH_2{-}CH_2$ (Phenyl, Phenyl)
	Styren $M_0 = 104$	*Distyren* $M = 208$
Polymerisationsgrad	1	2 ⩽ P ≪ 100

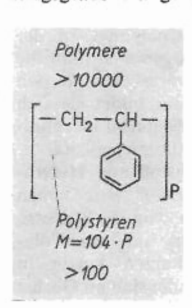


Abb. 5.4.3-1 Polymerhomologe des Styren

Im einfachsten Fall sind die Makromoleküle eines synthetischen Polymeren Glieder einer polymerhomologen Reihe. Sie stimmen hinsichtlich Struktur und Zusammensetzung überein, aber unterscheiden sich im Polymerisationsgrad. *Oligomere* sind die Anfangsglieder mit molaren Massen unter 10000 $gmol^{-1}$ (Abb. 5.4.3-1).
5.4.3.2. Polymerisationsgradverteilung. Die Polymolekularität wirkt sich auf viele Eigenschaften aus. So ergibt z. B. von zwei Polypropylenproben mit gleichem mittlerem Polymerisationsgrad die mit der engeren Verteilung Fasern mit weit höherer Zugfestigkeit.
5.4.3.2.1. *Verteilungsfunktionen.* Abbildungen von Polymerisationsgradverteilungen (↑ Abb. 5.4.3-3, 5.4.3-4) enthalten, genau genommen, keine stetigen Kurven, sondern Punktfolgen, da sich der Polymerisationsgrad immer nur um eine ganze Einheit ändern kann. Die Punkte liegen jedoch so dicht beisammen, daß praktisch Kurvenzüge resultieren. Diesen werden Funktionen zugeordnet, bei denen P als stetig veränderlich zwischen 0 und ∞ angesehen wird. Die *Häufigkeitsverteilungsfunktion* $n_P = h(P)$ gibt den in einem Grundmol der Polymersubstanz vorliegenden Anteil n_P in Mol an, dessen Polymerisationsgrad zwischen P und $P + 1$ liegt. Die *Massenverteilungsfunktion* gibt den in einem Gramm des Stoffs vorhandenen Massenanteil m_P in Gramm an, dessen Polymerisationsgrad zwischen P und $P + 1$ liegt. Die Massenverteilungsfunktion kann aus der Häufigkeitsverteilungsfunktion durch Multiplikation mit dem Polymerisationsgrad erhalten werden: $m_P = H(P) = Ph(P)$.
5.4.3.2.1.1. *Statistische Berechnung der Kettenlängenverteilung.* Die Form der Verteilungsfunktion wird durch die statistischen Gesetze bestimmt, denen die Aufbaureaktionen unterworfen sind. Für die Massenverteilung von Polymerisation läßt sich z. B. die Beziehung $m_P = H(P) = [c\Gamma(k + 1)]^{-1} (P/c)^k \exp(-P/c)$ ableiten, wobei ganzzahlige Werte des Koppelungsgrads k angenommen wurden. Die Gammafunktion $\Gamma(k + 1)$ steht für $\int_0^\infty t^k e^{-t}\, dt$.
Diese sogenannte *Schulzverteilung* eignet sich nach entsprechender Anpassung der Parameter k und c sehr gut zur Wiedergabe von experimentell ermittelten Verteilungen. Für $k = 1$ geht sie in die Normalverteilung des Polymerisationsgrads, die *Schulz-Flory-Verteilung*, über.

5.4.3.2.2. *Experimentelle Ermittlung der Polymerisationsgradverteilung durch Fraktionierung.* Die klassischen Fraktioniermethoden beruhen darauf, daß die Löslichkeit in einer polymerhomologen Reihe mit steigender Molmasse abnimmt.

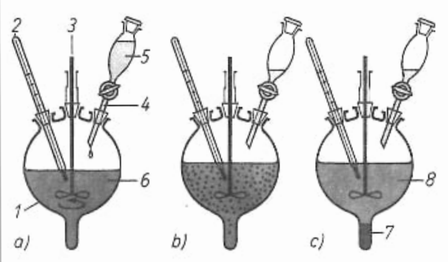


Abb. 5.4.3-2 Fraktionierte Fällung; a) Zutropfen des Fällungsmittels, b) Trübung durch ausgeschiedene Geltröpfchen, c) Absetzen der Gelphase beendet, Solphase nach Abtrennung zur nächsten Fällung bereit; *1* Fraktionierkolben mit Finger, *2* Thermometer mit Glasschliff, *3* Rührer mit Schliffhalterung, *4* Tropftrichter mit Schliff, Schliffhahn und Schliffstopfen, *5* Fällungsmittel, *6* Polymerlösung, *7* Gelphase, *8* Solphase

Fällungsfraktionierung. Gibt man zur Lösung eines makromolekularen Stoffs ein Fällungsmittel, das sich mit dem verwendeten Lösungsmittel mischt, so beginnt bei einer bestimmten Menge die Ausfällung der gelösten Substanz. Wird das Fällungsmittel so dosiert, daß zunächst nur ein kleiner Teil ausgeschieden wird, so kann man nach dem Absetzen des Niederschlags in der Gelphase (Abb. 5.4.3-2) den Anteil mit der höchsten Molmasse isolieren. Durch weitere, schrittweise Zusätze werden die Fraktionen mit niedrigeren Molmassen gewonnen. Das Polymer kann auch durch Abkühlen der Lösung gefällt werden.

Das Gegenstück dazu stellt die *Lösefraktionierung* dar, bei der die zu untersuchende Substanz in dünner Schicht auf einen Träger, z. B. einer Aluminiumfolie, aufgebracht wird. Dieser Filz wird mit Lösungsmitteln zunehmender Lösekraft extrahiert, die jeweils eine Fraktion herauslösen. Die Lösefraktionierung kann auf körnigen Trägern, z. B. Sand, auch in einer Säule und mit kontinuierlicher Änderung des Lösevermögens durchgeführt werden. Kombiniert man den Lösungsmittelgradient noch mit einem Temperaturgradient längs der Säule, so verwirklicht man das Prinzip der *Fällungschromatographie*, mit der zuerst automatisch sehr scharfe Fraktionen erhalten werden konnten. Diese Methode hat für präparative Fraktionierungen und für die Analyse der chemischen Uneinheitlichkeit in Copolymeren auch heute noch Bedeutung.

5.4.3.2.3. *Auswertung der Fraktionierergebnisse.* Aus dem Masseanteil und dem mittleren Polymerisationsgrad der einzelnen Fraktionen wird die *integrale Masseverteilungskurve* konstruiert, indem die Massebalken summierend über dem jeweiligen zugehörigen Polymerisationsgrad aufgetragen werden. Meist wird die so entstehende Treppenkurve durch eine, durch die Massemittelpunkte gelegte, stetige Kurve ersetzt (Abb. 5.4.3-3a). Da sich diese Integralkurve dem 100-%-Wert mit geringer Steigung nähert, besteht hinsichtlich des höchsten, im Polymeren vorkommenden Polymerisationsgrads eine beträchtliche Unsicherheit. Diese kann überwunden werden, wenn die Masseanteile auf einer nach dem Gaußschen Fehlerintegral geteilten Ordinate summierend über dem Logarithmus des Polymerisationsgrads aufgetragen werden (Abb. 5.4.3-3b). Durch graphische Differentiation der Integralkurve erhält man die *differentielle Masseverteilungskurve*, aus der entnommen werden kann, wie groß der Masseanteil m_P der Moleküle mit einem bestimmten Polymerisationsgrad ist (Abb. 5.4.3-4a). Sie gibt ein anschaulicheres Bild von der Zusammensetzung der Substanz als die Integralkurve (↑ Abb. 5.4.3-3a).

Die *Häufigkeitsverteilungskurve*, bei der die

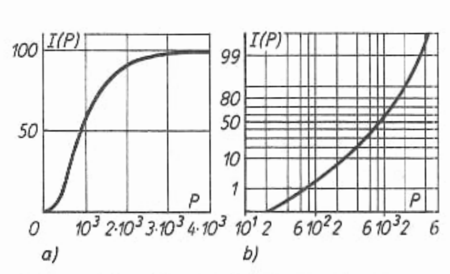


Abb. 5.4.3-3 Integrale Masseverteilungskurve eines Polymers mit Schulzverteilung; a) Darstellung im linearen Koordinatensystem, b) Wiedergabe im Summenwahrscheinlichkeitsnetz; *P* Polymerisationsgrad, *I*(*P*) Masseanteil bis zum Polymerisationsgrad *P* in %

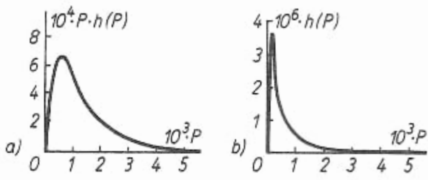


Abb. 5.4.3-4 a) Masseverteilung des in Abb. 5.4.3-3 dargestellten Polymers, b) zugehörige Häufigkeitsverteilung

relative Häufigkeit n_P der Moleküle eines bestimmten Polymerisationsgrads über P aufgetragen ist, ergibt sich aus der differentiellen Masseverteilungskurve durch Division der Ordinatenwerte aller Punkte durch den zugehörigen Polymerisationsgrad. Sie zeigt, wie groß die Zahl der Teilchen vom Polymerisationsgrad P in einem Mol der betrachteten Substanz ist (↑ Abb. 5.4.3-4b).

5.4.3.2.4. *Gelpermeationschromatographie.* Das moderne, abgekürzt mit *GPC* bezeichnete Verfahren trennt gelöste Moleküle in Säulen mit porösen Füllkörpern nach ihrer Größe. Sehr große Teilchen sind von den Poreninnenräumen ausgeschlossen und kommen zuerst am Säulenende an. Teilchen, die eindringen können, erscheinen später (Abb. 5.4.3-5). Zur GPC werden Füllkörper benötigt, deren Porendurchmesser den Abmessungen der zu trennenden Moleküle entsprechen. Für manche Trennaufgaben werden auch heute noch Gele (↑ 6.6.7.2.4.) eingesetzt, für andere Aufgaben verwendet man druckbeständige Füllkörper, z. B. poröses Glas, makroporös vernetzte Polymere (↑ 5.4.4.6.) oder Füllkörper auf Kieselsäurebasis. Die GPC weist eine Reihe von wesentlichen Vorteilen auf: Es wird wenig Substanz benötigt, die als Lösung (0,1 bis 0,3%) unmittelbar vor der Säule in den Elutionsmittelstrom eingespritzt wird. Die Probenanteile können mit Detektoren, die z. B. auf die Änderung des Brechungsindex Δn oder die UV-Absorption reagieren, kontinuierlich erfaßt werden. Die Auftrennung läßt sich mit Proben

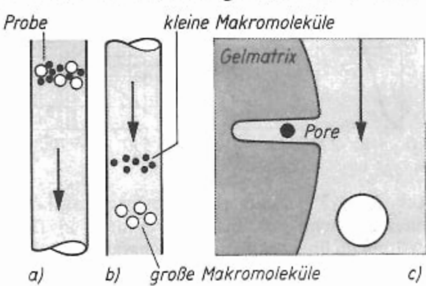


Abb. 5.4.3-5 Prinzip der GPC; a) Anfang, b) Ende der Trennsäule, c) Gelkern (schematisch)

von bekannter molarer Masse so eichen, daß aus dem Zeitpunkt der Elution eines bestimmten Probenanteils auf seine Molmasse geschlossen werden kann (Abb. 5.4.3-6a). In günstigen Fällen kann die Polymerisationsgradverteilung direkt aus der vom Detektor aufgezeichneten Kurve entnommen werden (Abb. 5.4.3-6b). Unter optimalen Bedingungen kann das Analysenergebnis 10 min nach der Injektion der Probe vorliegen.

Mittelwerte eignen sich als einfache Zahlenangaben gut für den orientierenden Vergleich von Proben. Um diesen Vorteil auch nach Ausführung von GPC-Analysen zu nutzen, werden aus den erhaltenen Verteilungskurven häufig die Mittelwerte und die Heterogenität nach den bekannten Summierungsvorschriften (↑ 5.4.3.3.) berechnet.

5.4.3.2.5. *Resultate.* Synthetische Polymere haben oft ziemlich breite Verteilungen, besonders wenn es sich um Abmischungen aus verschiedenen Chargen handelt oder wenn während der Polymerisation die Versuchsbedingungen wie Temperatur oder Initiatorkonzentration variiert wurden. Solche Maßnahmen können bei der großtechnischen Herstellung von Polymeren Vorteile bringen; sie verbessern den Umsetzungsgrad oder gewährleisten gleichmäßigen Produktionsausstoß ohne Kennzifferschwankungen. In extremen Fällen haben die Verteilungskurven sogar mehrere Maxima. Bei polymeren Naturstoffen werden solche Erscheinungen durch Wachstumsstörungen hervorgerufen.

Polymere mit enger Verteilung können entstehen, wenn bei einer Polymerisation alle Ketten zur gleichen Zeit gestartet werden und ihr Wachstum bis zum Verbrauch des Monomeren anhält. Dies ist z. B. bei der anionischen Polymerisation von α-Methyl-Styren in Tetrahydrofuran mit Naphthalen-Natrium als Initiator der Fall. Man spricht hier von *lebenden Polymeren*, da bei absolutem Ausschluß von Sauerstoff und Feuchtigkeit die Ketten ihre Fähigkeit, Monomereinheiten anzulagern, behalten können, bis die Abbruchreaktion durch Zusatz von Alkohol, Kohlendioxid oder dergleichen ausgelöst wird. Solange noch Monomeres vorhanden ist, hängt die erreichte

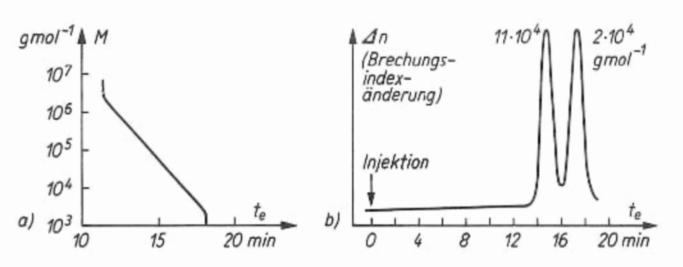


Abb. 5.4.3-6 a) GPC-Eichkurve: Zusammenhang zwischen Elutionszeit t_e und Molmasse, b) GPC-Elutionskurve: Gemisch aus zwei Polystyrenproben nach GPC-Trennung; Fließgeschwindigkeit 1 ml min^{-1}

Molekülgröße von der Dauer des Wachstums ab. Die Bildung von Makromolekülen einheitlicher Größe ist also in diesem Fall eine *zeitgesteuerte* Reaktion. Auch im lebenden Organismus können Polymere mit enger Verteilung entstehen. In Eiweißstoffen sind die Makromoleküle sogar völlig identisch. Das ist das Resultat der *strukturgesteuerten* Peptidsynthese.

5.4.3.3. Mittlerer Polymerisationsgrad. Dieser hat den Vorteil, daß Vergleiche besonders einfach möglich sind, doch ist ihre Aussagekraft begrenzt: Gleiche Mittelwerte können unter Umständen von verschiedenen Verteilungen und folglich von verschiedenen Produkten stammen.

Zum Mittelwert tragen alle in dem Produkt vorhandenen Anteile entsprechend ihrer Konzentration bei. Verschiedene Meßmethoden können unterschiedliche Mittelwerte ergeben, wenn z. B. eine Meßgröße den einzelnen Beiträgen P_i direkt proportional und eine andere Meßgröße dem Polymerisationsgrad der Anteile umgekehrt proportional ist. Für den letztgenannten Fall läßt sich $\bar{P}_n = (\sum n_i P_i)/(\sum n_i)$ ableiten. $\bar{P}_n$ ist der *Zahlenmittelwert* [engl. *number average*] des Polymerisationsgrads. Der *Gewichtsmittelwert* P_w [engl. *weight average*] ergibt sich bei direkter Proportionalität zu $\bar{P}_w = (\sum n_i P_i^2)/(\sum n_i P_i)$. Bei der Bestimmung der Molmasse aus der Viskosität $[\eta]$ verdünnter Lösungen gilt $[\eta] = KP^a$. Derartige Messungen liefern $\bar{P}_v^a = (\sum n_i P_i^{a+1})/(\sum n_i P_i)$ für das *Viskositätsmittel*. Die in diesen Formeln durchzuführenden Summierungen erfassen alle vorhandenen Werte des Polymerisationsgrads; n_i ist die Molzahl des Anteils mit dem Polymerisationsgrad P_i, a ist eine Konstante.

Das Viskositätsmittel wird mit dem Gewichtsmittel identisch, wenn $a = 1$ ist. In der Regel liegt der Wert für a bei 0,8, daher ist $\bar{P}_w$ im allgemeinen größer als $\bar{P}_v$.

Mit dem Verhältnis $\bar{P}_w/\bar{P}_n$ kann die *Heterogenität* eines Polymers charakterisiert werden. Liegt eine Schulzverteilung vor (↑ 5.4.3.2.1.), so ist $\bar{P}_n/k = \bar{P}_w(k+1)$. Daraus resultiert für die Normalverteilung $(k = 1)$ $\bar{P}_w = 2\bar{P}_n$; zugleich ist $\bar{P}_n$ das Maximum der Masseverteilungskurve (Abb. 5.4.3-7).

Wegen des Zusammenhangs $M = PM_0$ (↑ 5.4.2.) können analoge Betrachtungen und Berechnungen auch für die Mittelwerte der molaren Masse durchgeführt werden.

5.4.3.4. Absolutmethoden zur Bestimmung des mittleren Polymerisationsgrads bzw. der mittleren molaren Masse. Die Absolutmethoden liefern die Molmasse als Funktion von meßbaren Größen und von Faktoren, die aufgrund der Theorie des Bestimmungsverfahrens berechnet werden können. Sie bedürfen daher im Unterschied zu den Relativmethoden keiner Eichung mit Proben von bekannter Molmasse.

5.4.3.4.1. *Osmometrie.* Geräte zur Messung des osmotischen Drucks (↑ 6.3.3.3.5.) enthalten eine Membran, die für das Lösungsmittel durchlässig und für die Moleküle des Gelösten undurchlässig ist. Der sich aufgrund der Tendenz zum Konzentrationsausgleich einstellende osmotische Druck ist dem Polymerisationsgrad

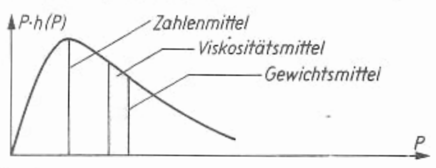


Abb. 5.4.3-7 Masseverteilungsfunktion (schematisch) mit Angabe der Mittelwerte

umgekehrt proportional und ergibt für polymolekulare Proben einen Zahlenmittelwert (↑ 5.4.3.3.).

5.4.3.4.2. *Ebullioskopie.* Die Dampfdruckerniedrigung (↑ 6.3.3.3.2.) von Polymerlösungen beträgt nur etwa 10^{-1} Pa, d. h. 10^{-4}% des Ausgangswerts und läßt sich nur sehr schlecht direkt messen. Die mit ihr einhergehende Siedepunkterhöhung (↑ 6.3.3.3.4.) kann jedoch zur Bestimmung von Molmassen bis zu 25000 g mol^{-1} dienen. Dabei müssen Temperaturänderungen in der Größenordnung 10^{-4} bis 10^{-5} K gemessen werden. Um den Einfluß wetterbedingter Schwankungen des Luftdrucks auszuschalten, verwendet man am besten eine Apparatur mit zwei gleichartigen Siedegefäßen, in denen Lösung und Lösungsmittel zur gleichen Zeit unter identischen Bedingungen untersucht werden können.

5.4.3.4.3. *Isotherme Destillation.* Hält man zwei Lösungen verschiedener Konzentration in einem gemeinsamen Raum auf gleicher Temperatur, so bewirkt der Dampfdruckunterschied eine Lösungsmitteldestillation. Dabei nimmt die Lösungsmittelmenge der konzentrierteren Lösung zu, die der verdünnteren ab. Durch Vergleich mit Lösungen bekannter Konzentrationen läßt sich die unbekannte Teilchenzahl in der Untersuchungslösung eingrenzen. Die Methode ist zwar zeitaufwendig, erfordert aber nur geringe apparative Mittel. Sie kann unter dem Mikroskop in beiderseits zugeschmolzenen Schmelzpunktkapillaren mit sehr kleinen Materialmengen durchgeführt werden.

Ebullioskopie und isotherme Destillation sprechen auf die Anzahl der in einem bestimmten Volumen vorhandenen Moleküle an. Da diese bei gegebener Masse dem Polymerisationsgrad umgekehrt proportional ist, werden *Zahlenmittelwerte* bestimmt.

5.4.3.4.4. *Endgruppenbestimmung.* Manche Polymere haben Endgruppen, z. B. —COOH in

Polyestern oder Polyamiden, die sich durch *Titration*, durch physikalische Messungen wie *IR-Absorption* und *Kolorimetrie* oder durch *radiochemische Analysenmethoden* quantitativ bestimmen lassen. Wenn bekannt ist, wieviel solche Endgruppen in einem Makromolekül vorhanden sind, kann aus dem Ergebnis der Endgruppenbestimmung die Anzahl der in der vorgelegten Probe enthaltenen Moleküle und

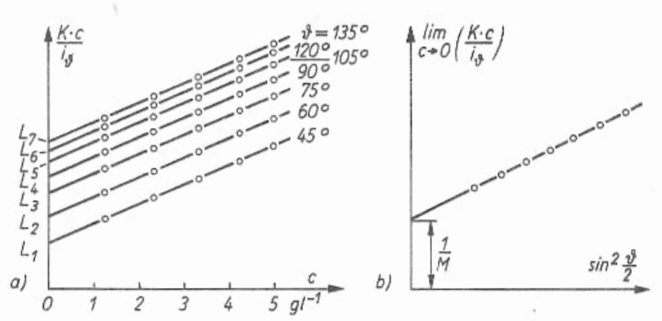


Abb. 5.4.3-8 a) Extrapolation der durch Lichtstreuung an großen Makromolekülen unter verschiedenen Winkeln ϑ gemessenen Werte auf die Konzentration $c = 0$, b) Extrapolation der aus a) entnommenen Ordinatenabschnitte auf den Winkel $\vartheta = 0$

damit der Zahlenmittelwert der Molmasse bestimmt werden. Mit chemischen Mitteln lassen sich Werte bis zu 25000 g mol^{-1} messen; bei den empfindlicheren radiochemischen und spektroskopischen Methoden liegt die Grenze bei 100000 g mol^{-1}.

5.4.3.4.5. *Lichtstreuung.* Die Teilchen lichtdurchlässiger Stoffe streuen einen Teil des einfallenden Lichts in seitliche Richtungen. DEBYE hat gezeigt, wie die von gelösten Makromolekülen ausgehende Streuintensität i_{red} mit der molaren Masse M zusammenhängt: Bei linear polarisiertem Licht und Beobachtung in der zur Polarisationsrichtung senkrechten Ebene gilt für unendlich kleine Konzentrationen c als Grenzgesetz $i_{red} = KcM$. Die Konstante K enthält unter anderem die Wellenlänge λ des für die Anregung benutzten Lichts und das Brechungsinkrement $(\partial n/\partial c)$. Da Messungen nur bei endlichen Konzentrationen möglich sind, muß eine Gleichung der Form $(Kc)/i_{red} = 1/M + 2A_2c$ benutzt werden, in der mit A_2 der 2. Virialkoeffizient auftritt und die für $c \to 0$ in die Grenzformel $i_{red} = KcM$ übergeht. Gemessen wird i_{red} bei der Temperatur T und verschiedenen Konzentrationen c. Da K bekannt ist, kann der Quotient Kc/i_{red} gebildet und über c aufgetragen werden. Alle Kurven für dieselbe Substanz in verschiedenen Lösungsmitteln schneiden die Ordinate beim Wert $1/M$. Aus der Steigung der geraden Anfangsabschnitte lassen sich die Virialkoeffizienten A_2 berechnen, die mit der thermodynamischen Güte des Lösungsmittels zunehmen.

Wenn die Abmessungen der Makromoleküle größer sind als $\lambda/20$, dann tritt zwischen den an verschiedenen Teilen eines Makromoleküls gestreuten Lichtwellen *Interferenz* ein. Diese führt zu einer Intensitätsminderung, die um so ausgeprägter ist, je größer der Winkel ϑ zwischen Primärstrahl und Streurichtung und je größer das streuende Teilchen ist. Die Lichtstreuung hängt dann vom Beobachtungswinkel ab. In diesem Fall kann nach der Beziehung $(Kc)/i_{red} = 1/M + 2A_2c + k(d^2/\lambda^2)\sin^2(\vartheta/2)$ ausgewertet werden, in der k eine Konstante, d die Molekülgröße, λ die Wellenlänge und ϑ der Beobachtungswinkel sind. Die Gleichung zeigt, daß $Kc/i = 1/M$ ist, wenn $\vartheta = 0$ und $c = 0$ wird. Diese beiden Grenzübergänge können nacheinander ausgeführt werden. Trägt man die Meßwerte zuerst über c auf (Abb. 5.4.3-8a), so bestimmen die zu gleichen Winkeln gehörenden Punkte Geraden mit den Ordinatenabschnitten $\lim\limits_{c\to 0}(Kc/i_{red})$. Diese ergeben über $\sin^2(\vartheta/2)$ abermals eine Gerade (Abb. 5.4.3-8b) mit dem Ordinatenabschnitt $1/M$. Neben der Molmasse ergibt diese Darstellung mit dem Anstieg der Geraden die Größe d, die die Abmessung der streuenden Teilchen charakterisiert (s. a. 5.4.6.1.). Bei umgekehrter Reihenfolge der Grenzübergänge wird in analoger Weise der Virialkoeffizient A_2 zugängig. Beide Wege erfordern vier Einzeldiagramme. Doch lassen sich alle Extrapolationen in einer einzigen Darstellung, dem *Zimm-diagramm* (Abb. 5.4.3-9), vornehmen, wenn Kc/i_{red} über der Summe $\sin^2(\vartheta/2) + \alpha c$ aufgetragen wird. Die Größe α hat die Dimension einer reziproken Konzentration. Ihr Zahlenwert wird so gewählt, daß die Meßpunkte eine Figur mit nicht zu steilen Flanken ergeben. Meist ist $\alpha = 100$.

Bei dieser Darstellung bestimmen die Werte, die an der Lösung mit der höchsten Konzentration unter den verschiedenen Winkeln gemessen worden sind, im Idealfall eine Gerade, die die obere Seite eines Parallelogramms bildet. Parallel dazu verlaufen die an den Lösungen geringerer Konzentration bestimmten Kurven. Diese Schar von Geraden wird nach links bis zum Schnittpunkt mit der über dem jeweiligen

Wert αc errichteten Senkrechten verlängert. Die Schnittpunkte stellen die Extrapolation nach dem Winkel $\vartheta = 0$ dar. Miteinander verbunden ergeben sie eine Gerade, die auf der Ordinate den Abschnitt $1/M$ markiert. Ihre Steigung liefert den Virialkoeffizienten A_2.
Werden dagegen die Punkte miteinander verknüpft, die an den verschiedenen Lösungen jeweils beim größten Winkel gemessen wurden, so ergibt sich die steile rechte Begrenzungslinie des Parallelogramms. Parallel dazu, aber nach links versetzt, liegen die Geraden, die aus den Meßwerten bei niedrigeren Winkeln resultieren. Diese Geraden werden nach unten bis zu den zugehörigen Abszissenwerten $\sin^2(\vartheta/2)$ verlängert, die von der auf der Abszisse aufgetragenen Summe übrig bleiben, wenn $c = 0$ wird. Die Schnittpunkte mit den bei diesen Werten gezogenen Senkrechten stellen also die Extrapolation $c \to 0$ dar. Sie bestimmen eine Gerade, die die Ordinate beim gleichen Wert $1/M$ teilt wie die Extrapolationsgerade für $\vartheta = 0$, und aus deren Neigung die *Moleküldimensionen* berechnet werden können.

5.4.3.4.6. *Sedimentation in der Ultrazentrifuge.* Teilchen der Masse m setzen sich unter der Wirkung einer Beschleunigung b mit einer Geschwindigkeit ab, die von dem Produkt mb bestimmt wird. Während Sand- und Tonsuspensionen schon unter der Wirkung der Erdbeschleunigung g sedimentieren, erfordern Moleküle größere Beschleunigungen, die in hochtourigen Zentrifugen erzeugt werden. Mit der Ultrazentrifuge werden bei 1000 Hz (60000 $\min^{-1}$) Werte von mehr als $10^5 g$ erreicht. Der Rotor läuft in einer Wasserstoffatmosphäre mit vermindertem Druck. Optische Vorrichtungen ermöglichen die Beobachtung der Küvette im Rotor bei vollem Lauf (Abb. 5.4.3-10).
Bei sehr hohen Werten der Zentrifugalbeschleunigung wandern die gelösten Makromoleküle bis zum Boden der Küvette. Ihre Geschwindigkeit wird durch die antreibende Sedimentationskraft und die bremsende Reibungskraft bestimmt und ist bei kugelförmigen Partikeln eine eindeutige Funktion der Teilchenmasse.
Da mit dieser Funktion auch die molare Masse von Kettenmolekülen richtig ermittelt werden kann, muß angenommen werden, daß diese nicht als Stäbchen oder vom Lösungsmittel frei durchspülte „Drahtknäuel“ in der Lösung vorliegen können. Sie sedimentieren vielmehr gemeinsam mit einer gewissen Menge Lösungsmittel, das im Knäuel *immobilisiert* ist. Die Molmasse wird mit der Svedberggleichung $M = (RTs)/[D(1 - v\varrho)]$ berechnet, in der v das spezifische Volumen des Gelösten und ϱ die Dichte des Lösungsmittels bedeuten. Der Ausdruck $(1 - v\varrho)$ korrigiert den Auftrieb durch das verdrängte Lösungsmittel. *Diffusionskoeffizient D* und *Sedimentationskoeffizient s*, d. h. das Verhältnis der beobachteten Geschwindigkeit zur wirkenden Beschleunigung, sind konzentrationsabhängig; es müssen Messungen bei verschiedenen Konzentrationen durchgeführt werden, die zur Konzentration Null zu extrapolieren sind.
Bei kleineren Werten des Produkts mb setzen sich die Teilchen nicht vollständig ab, weil die *Brownsche Molekularbewegung* der Sedimentation entgegenwirkt. Es bildet sich eine Verteilung im Zentrifugalfeld aus, die mit der baro-

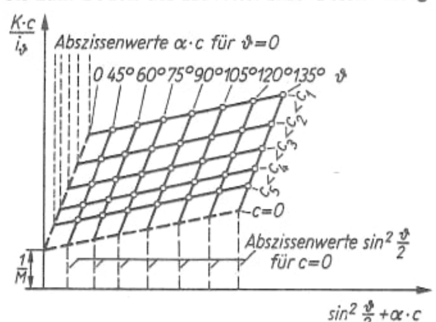


Abb. 5.4.3-9 Auswertung der an großen Makromolekülen gemessenen Lichtstreuung im Zimmdiagramm

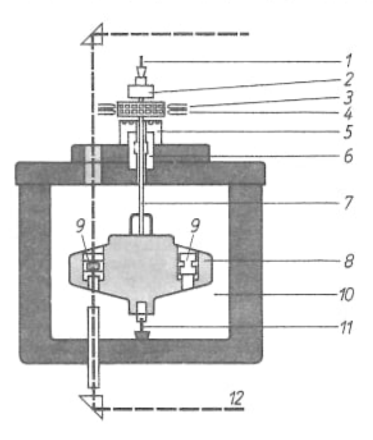


Abb. 5.4.3-10 Schematischer Querschnitt einer Ultrazentrifuge; *1* oberer und *11* unterer Kontakt für die elektrische Messung der Temperatur im Rotor, *2* Geber für elektronische Drehzahlregelung, *3* Treibdüse, *4* Bremsdüse, *5* Hubluftdüse, *6* Gleitlager, *7* Achse, *8* Rotor, *9 links* Meßzelle und *rechts* Ausgleichskörper, *10* Vakuumkammer, *12* Strahlengang für die Beobachtung der Sedimentation während des Laufs der Zentrifuge

metrischen Höhenformel ausgewertet werden kann. Allerdings kann es Tage dauern, bis sich das Gleichgewicht eingestellt hat. Während dieser Zeit muß die Ultrazentrifuge ohne Temperatur- oder Drehzahlschwankungen in Betrieb gehalten werden.

5.4.3.5. Relativmethoden zur Bestimmung des mittleren Polymerisationsgrads. Im Unterschied zu den Absolutmethoden (↑ 5.4.3.4.) müssen die Relativmethoden mit Proben geeicht werden, von denen der Polymerisationsgrad genau bekannt ist. Er kann bei monodispersen Polymeren mit jeder beliebigen Absolutmethode bestimmt werden. Haben die Eichproben dagegen eine breite Verteilung, dann ist ein Absolutverfahren zu wählen, das den gleichen Mittelwert liefert wie die Methode, die geeicht werden soll. Erhält man bei der Relativmessung z. B. ein Gewichtsmittel, dann ist die Eichung aufgrund von Messungen vorzunehmen, die ebenfalls das Gewichtsmittel ergeben. Würde man die Relativmethode nach Zahlenmittelwerten eichen, so liefert sie nur dann richtige Ergebnisse, wenn die Verteilung in den unbekannten Proben zufällig mit der in den Eichsubstanzen übereinstimmt.

5.4.3.5.1. *Viskosimetrie.* Sie beinhaltet die Bestimmung der *spezifischen Viskosität* $\eta_{sp} = (\eta - \eta_0)/\eta_0$, die von der Viskosität η_0 für das reine Lösungsmittel abhängt. Der *Staudingerindex* $[\eta]$ ergibt sich als Grenzwert von (η_{sp}/c) für $c \to 0$ und ein ebenfalls gegen Null gehendes Geschwindigkeitsgefälle. Zum Polymerisationsgrad P besteht der Zusammenhang $[\eta] = KP^a$ (↑ 5.4.3.3.). Die erhaltenen Werte sind Viskositätsmittel, die dicht unter den Gewichtsmittelwerten liegen. Die Eichung der Viskosimetrie, d. h. die empirische Ermittlung der Konstanten K und a, erfolgt am besten mit Proben, deren Polymerisationsgrad durch Lichtstreuungs- oder Sedimentationsmessungen bestimmt wurde.

Wenn eine Eichung nach absoluten Molmassen nicht möglich ist, kann die Viskosimetrie nur zur Unterscheidung von Polymerproben herangezogen werden. Dazu kann der Staudingerindex, die spezifische Viskosität bei einer bestimmten Konzentration oder der in der Praxis immer noch benutzte *K-Wert* angegeben werden.

5.4.3.5.2. *Schmelzindex.* Bei Kunststoffen wie Polyethylen oder Polypropylen, die bei Zimmertemperatur unlöslich sind, wird zur Charakterisierung der Molekülgröße oft die Schmelzviskosität herangezogen. Sie steigt mit wachsender Molmasse stark an, hängt aber auch von der Breite der Verteilung ab. Der Schmelzindex gibt an, wieviel Gramm Polymerschmelze in einer bestimmten Zeit aus einer Düse austreten, wenn bestimmte Bedingungen hinsichtlich Druck, Temperatur und Düsengeometrie eingehalten werden. Je dünner die Schmelze ist, um so mehr wird ausgepreßt; der Schmelzindex ist daher um so größer, je niedriger der Polymerisationsgrad liegt.

5.4.3.5.3. *Fällungstitration.* Fügt man zu einer makromolekularen Lösung einen mit dem Lösungsmittel mischbaren Nichtlöser, so beginnt bei einer bestimmten Menge die Ausfällung der gelösten Substanz. Je niedriger der Polymerisationsgrad ist, um so mehr Fällungsmittel wird benötigt. Für äußerst verdünnte Lösungen gilt für den Volumenbruch am Trübungspunkt $\varphi^* = A + B(P^{-0,5})$, sofern Temperatur und Polymerkonzentration konstant gehalten werden. Die empirischen Konstanten A und B sind durch Eichung zu ermitteln. Der Trübungspunkt ist bei Proben mit enger Verteilung sehr gut zu erkennen. In diesem Fall ist die Ermittlung des Polymerisationsgrads aus dem Fällungsmittelverbrauch mit genügender Sicherheit möglich. Die Fehler gehen bis zu etwa $P = 1000$ nicht über die bei anderen Methoden gegebenen Grenzen hinaus. Die Fällungstitration ermöglicht, den Polymerisationsgrad mit einfachen Mitteln und geringem Zeitaufwand an sehr kleinen Proben in günstigen Fällen mit weniger als 0,5 mg Substanz zu bestimmen.

5.4.3.5.4. *Thermoelektrische Messung der Verdampfungs- oder Kondensationsgeschwindigkeit (Hill-Baldes-Methode).* Die Dampfdruckerniedrigung bewirkt, daß eine Lösung langsamer verdampft als das reine Lösungsmittel (s. a. 5.4.3.4.3.). Da sich eine verdampfende Flüssigkeit abkühlt, kann die unterschiedliche Verdampfungsgeschwindigkeit zu einer Temperaturdifferenz zwischen Lösung und Lösungsmittel führen. Dieser Effekt wird mit dem Dampfdruckosmometer gemessen. In ihm befinden sich zwei gleichartige, hochempfindliche Temperaturfühler in einem temperaturkonstanten Raum, der Dampf des verwendeten Lösungsmittels enthält. An der einen Meßstelle hängt ein Tropfen der zu untersuchenden Lösung, an der anderen ein Tropfen des Lösungsmittels. Aus der unter definierten Bedingungen gemessenen Temperaturdifferenz können Molmassen bis zu 40000 $g\,mol^{-1}$ und mit wenigen Milligramm Substanz bestimmt werden.

5.4.4. Konstitution der Makromoleküle

Am Aufbau der Makromoleküle ist in der Regel Kohlenstoff oder ein anderes vier- oder dreiwertiges Element maßgeblich beteiligt. Wenn diese mehrwertigen Elemente Makromoleküle ergeben, deren Gerüst sich nur in einer Richtung erstreckt, sind alle nicht in der Kette betätigten Valenzen mit Atomen wie

Wasserstoff oder Halogen oder mit kleinen Atomgruppierungen wie $—CH_3$ oder $—OH$ abgesättigt.

Die mehrwertigen Atome können aber auch zweidimensionale oder räumliche Netzwerke bilden, wenn sie mit den Seitenvalenzen direkt oder über Zwischenglieder zusätzlich miteinander verknüpft sind. Flächige Strukturen sind selten. Daher beschränken sich die folgenden Darstellungen auf kettenförmige und auf dreidimensionale Polymere.

5.4.4.1. Lineare Homopolymere. Homopolymere sind die makromolekularen Reaktionsprodukte von Aufbaureaktionen, in die nur eine Art von Monomeren eingeht. Sie sind *regelmäßige Polymere*, wenn ihre Struktureinheiten alle dieselbe sequentielle Anordnung haben. Für Vinylpolymere kommen drei Möglichkeiten in Betracht (Abb. 5.4.4-1). Die regelmäßige 1,2-Struktur überwiegt; hohe Polymerisationstemperaturen führen jedoch zur Störung der Regelmäßigkeit durch einzelne 1,1-Anlagerungen, vor allem wenn der Substituent an C(1) nur geringe dirigierende Wirkung entfaltet. Ringförmige Monomere wie Caprolactam (Rohstoff für Polyamid 6) oder Ethylenoxid ergeben regelmäßige Polymere (Abb. 5.4.4-2).

5.4.4.1.1. *Polydiene.* Bei der Polymerisation von Dienen, das sind Verbindungen mit zwei ungesättigten Doppelbindungen im Molekül, z. B. Butadien $CH_2{=}CH—CH{=}CH_2$, kommen zu den in Abb. 5.4.4-1 dargestellten Konstitutionsvarianten noch weitere hinzu. Es können 1,2-Polymerisate mit Vinylseitengruppen oder 1,4-Polymerisate mit Doppelbindungen in der Kette entstehen (Abb. 5.4.4-3). Sind die beiden Doppelbindungen im Monomeren durch eine biegsame Gruppierung *R* isoliert, so ist *Cyclopolymerisation* möglich (Abb. 5.4.4-4). Polymere, in denen die Monomerstruktur so stark verändert auftritt, werden *Phantompolymere* genannt.

5.4.4.1.2. *Endgruppen.* Die Wiedergabe von Ausschnitten aus hochpolymeren Ketten bricht beiderseits mit der Andeutung der Valenzen ab, die zu den nächsten Grundbausteinen führen. Am Kettenende müssen diese Bindungen durch einwertige Gruppen abgesättigt werden, die ganz andere Struktur und Zusammensetzung als die Grundbausteine haben können. Bei Polymerisation können es z. B. Bruchstücke der Initiatoren sein, die die Startreaktion ausgelöst haben. Die Endgruppen können aber auch von Überträgersubstanzen oder von kettenabbrechenden Stoffen herrühren. Andererseits können sie in enger Verwandtschaft zu den Kettenbausteinen stehen. So befindet sich in unstabilisierten Polykondensaten einfach eine Monomereinheit mit einer freien funktionellen Gruppe am Kettenende. Bei Polymerisaten können nach Übertragungsreaktionen durch das Monomer oder Kettenabbruch durch Disproportionierung die Gruppierungen $—CH_2R$, $CH_2{=}CR—$ oder $—CH{=}CHR$ die Ketten begrenzen. Obgleich die Endgruppen von den

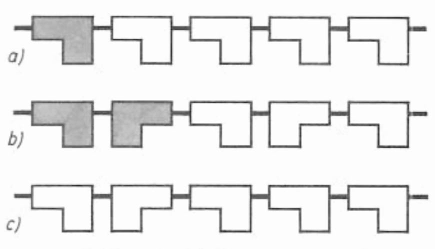


Abb. 5.4.4-1 Konstitution von Vinylpolymeren; a) regelmäßige 1,2-Verknüpfung der Monomereinheiten, b) 1,1-Verknüpfung (Kopf-Kopf-Struktur), c) unregelmäßige Verknüpfung

a) $-CH_2-CH_2-O-CH_2-CH_2-O-CH_2-CH_2-O-$

b) $-N(H)-(CH_2)_5-C(=O)-N(H)-(CH_2)_5-C(=O)-N(H)-(CH_2)_5-C(=O)-$

Abb. 5.4.4-2 Ausschnitt aus der Kette a) von Polyethylenoxid und b) von Polyamid 6

a) $-\overset{1}{C}H_2-\overset{2}{C}H(-\overset{3}{C}H{=}\overset{4}{C}H_2)-\overset{1}{C}H_2-\overset{2}{C}H(-\overset{3}{C}H{=}\overset{4}{C}H_2)-\overset{1}{C}H_2-\overset{2}{C}H(-\overset{3}{C}H{=}\overset{4}{C}H_2)-$

b) $-\underset{1}{C}H_2-\underset{2}{C}H{=}\underset{3}{C}H-\underset{4}{C}H_2-\underset{1}{C}H_2-\underset{2}{C}H{=}\underset{3}{C}H-\underset{4}{C}H_2-$

Abb. 5.4.4-3 a) 1,2-Polybutadien, b) 1,4-Polybutadien

$$n\,\overset{1}{C}H_2{=}\overset{2}{C}H-\overset{3,4,5}{(CH_2)_3}-\overset{6}{C}H{=}\overset{7}{C}H_2$$

$$\downarrow$$

$$-\overset{1}{C}H_2-\overset{2}{C}H-\overset{6}{C}H-\overset{7}{C}H_2-\overset{1}{C}H_2-\overset{2}{C}H-\overset{6}{C}H-\overset{7}{C}H_2-$$

(ring in each unit: $\overset{2}{C}H-\overset{3}{C}H_2-\overset{4}{C}H_2-\overset{5}{C}H_2-\overset{6}{C}H$)

Abb. 5.4.4-4 Cyclopolymerisation von α,ω-Heptadien

Grundbausteinen verschieden sind, werden sie bei den üblichen chemischen Analysen in der Regel nicht bemerkt, weil ihre Konzentration sehr gering ist. In einem Polymer mit dem Polymerisationsgrad $P = 1000$ befinden sich z. B. nur 0,2% Endgruppen, d. h. 2 auf 1000 Monomereinheiten. Durch gezielte Analysen hoher Empfindlichkeit können sie jedoch qualitativ und quantitativ bestimmt werden (s. a. 5.4.3.4.4.).

5.4.4.2. Lineare Copolymere. Copolymere sind Polymere aus mehr als einer Art von Monomeren. Je nach Zahl der beteiligten Arten unterscheidet man *Bipolymere, Terpolymere* oder *Quaterpolymere.* In *statistischen Copolymeren* sind die verschiedenen Monomerarten statistisch verteilt; in *Blockcopolymeren* bestehen größere Abschnitte der Polymerkette nur aus einem Monomeren. *Diblock-* oder *AB*-Blockcopolymere sind Makromoleküle, in denen eine Homokette aus dem Monomer *A* endständig mit einer Homokette *B* verknüpft ist. Ein *ABA*-Blockcopolymer ist ein *Triblockcopolymer* aus den beiden Monomeren *A* und *B*, es besitzt zwei *A*-Sequenzen und eine *B*-Sequenz. Die *Sequenzlänge* ist die Anzahl der sich wiederholenden Einheiten in einem nur aus einer Monomerart bestehenden Molekülabschnitt. Mittelwerte der Sequenzlänge werden für jede Monomerart analog wie die Mittelwerte des Polymerisationsgrads definiert (↑ 5.4.3.3.). In statistischen Copolymeren ist die mittlere Sequenzlänge viel kleiner als in Blockcopolymeren, aber auf jeden Fall größer als Eins. Wird sie für beide Monomere *A* und *B* genau gleich Eins, so folgt immer eine *A*-Einheit auf eine *B*-Einheit und umgekehrt; es liegt ein *alternierendes Copolymer* vor. *Taperpolymere* [*taper* engl., Wachsstock] weisen von einem Kettenende zum anderen stetig abnehmende Sequenzlänge auf (Abb. 5.4.4-5). Die *Blockzahl* ist die Gesamtzahl aller Sequenzen in einem Kettenabschnitt mit 100 Monomerbausteinen.

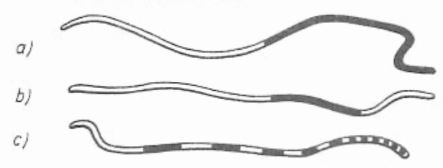


Abb. 5.4.4-5 Blockcopolymere; a) *AB*-Diblockcopolymer, b) *ABA*-Triblockcopolymer, c) Taperpolymer

Copolymere können auch durch Polykondensation oder Polyaddition entstehen, wenn von Gemischen von Monomeren ausgegangen wird, die sich in der Kette gegenseitig vertreten können. So werden ungesättigte Polyester aus Diolen und Gemischen von Dicarbonsäuren hergestellt, die z. B. Phthalsäure- neben Maleinsäureanhydrid enthalten. Diese Säurereste treten entsprechend im Polyestermolekül auf, wo sie jeweils durch Glycoleinheiten verknüpft sind.

Ein besonders interessantes Problem ist die Anordnung von austauschbaren Bausteinen zu bestimmten Sequenzen, wie es bei den Eiweißstoffen der Fall ist, die aus Aminosäuren aufgebaut sind. Wie bei den synthetischen Polyamiden sind die Grundbausteine durch die Bindung —CONH— miteinander verknüpft. Da die Proteine und Enzyme durch *strukturgesteuerte Aufbaureaktionen* (↑ 5.4.3.2.5.) entstehen, haben alle Ketten die gleiche Monomerfolge. Während die prozentuale Beteiligung der einzelnen Aminosäuren nach Totalhydrolyse des Eiweißstoffs verhältnismäßig einfach ermittelt werden kann, erfordert die Sequenzanalyse schrittweisen Abbau mit Isolierung und Identifizierung der bei jedem Schritt abgespaltenen Aminosäure. Die ersten Untersuchungen dieser Art am Insulin wurden mit bewundernswürdigem Fleiß manuell bewältigt. Jetzt gibt es Analysengeräte, die alle erforderlichen Schritte automatisch ausführen. Die zu untersuchenden Eiweißmoleküle werden mit einem Ende an einem festen Träger fixiert. Am anderen Ende wird durch chemische Reaktionen ein Baustein nach dem anderen abgekoppelt. Das muß bei allen 10^{10} Polymerketten, die für etwa 10 pg Analysensubstanz erforderlich sind, genau im gleichen Takt erfolgen. Die Sequenzanalysen haben bestätigt, daß jedes Eiweiß anders gebaut ist. Die etwa 20 verschiedenen Aminosäuren, die in den Proteinen vorkommen, können in so mannigfaltiger Weise zu Makromolekülen kombiniert werden, daß jedes Lebewesen seinen individuellen Eiweißstoff besitzen könnte.

Abgesehen von den eben besprochenen, durch strukturgesteuerte Aufbaureaktionen erzeugten Biopolymeren und von strikt alternierend gebauten Produkten, sind Copolymere hinsichtlich der *chemischen Zusammensetzung* immer *heterogen.* Das Monomerverhältnis weist eine *statistische Streuung um den Mittelwert* auf, die durch lokale Konzentrationsschwankungen in der unmittelbaren Nähe des Wachstumszentrums während der Polymerisation entsteht. Die statistische Streuung tritt folglich auch dann auf, wenn die mittlere Zusammensetzung des Copolymers in allen Phasen der Polymerisation mit der Bruttozusammensetzung des Monomergemisches übereinstimmt. Im allgemeinen wird jedoch eine Komponente bevorzugt eingebaut, an der das Ausgangsgemisch folglich verarmt. Dadurch entsteht zusätzlich die *umsatzbedingte Heterogenität.* Wenn diese unterdrückt werden soll, muß die Reaktion schon nach wenigen Prozent Umsatz

abgestoppt werden. Schließlich trägt auch die *sequentielle Heterogenität* zur chemischen Uneinheitlichkeit bei. Wenn man noch die stets vorhandene Heterogenität des Polymerisationsgrads (↑ 5.4.3.3.) berücksichtigt, so kann man ohne zu übertreiben behaupten, daß in einem Copolymer kaum zwei völlig gleiche Makromoleküle vorkommen.

Abb. 5.4.4-6 Aus Polyacrylnitril (a) entsteht durch thermische Cyclisierung (b) und Dehydrierung ein Leiterpolymer (c)

5.4.4.3. Doppelstrangpolymere. Diese Bezeichnung rührt von der Strickleiterstruktur her, die die Makromoleküle dieser Stoffe besitzen. Sie kann durch spezielle Behandlung von bestimmten Ausgangspolymeren realisiert werden. Beispielsweise entsteht aus Polyacrylnitril durch thermische Cyclisierung und nachfolgende Dehydrierung ein Leiterpolymer (Abb. 5.4.4-6), das bei hohen Temperaturen verwendet werden kann. Die Leiterpolymere haben wegen der vielen π-Elektronen Eigenschaften elektrischer Halbleiter und sind schwach gefärbt.

5.4.4.4. Verzweigte Polymere. Verzweigte Homopolymere entstehen gezielt durch Aufbaureaktionen mit Beteiligung von tri- oder polyfunktionellen Monomeren. Definiert verzweigte Polymere mit Stern- oder Kammstruktur werden aus lebenden Polymeren (↑ 5.4.3.2.5.) und entsprechenden multifunktionellen Initiatoren oder Kettenabbrechern hergestellt. Verzweigungen entstehen auch spontan beim Kettenaufbau, z. B. durch Übertragung des angeregten Zustands auf bereits vorliegende Polymermoleküle. Da Übertragungsreaktionen bei vielen Polymerisationen zum normalen Ablauf gehören, sind streng lineare Polymerisate selten, besonders wenn bis zu hohen Umsatzgraden polymerisiert wird. Chemische, fotochemische oder mechanische Aktivierung von Polymeren führt ebenfalls zu Verzweigungen. Im allgemeinen sind sowohl die Verteilung der Verzweigungsstellen über das Molekül als auch die Länge der entstehenden Seitenketten statistische Größen.

5.4.4.5. Pfropfpolymere. Die Makromoleküle dieser Produkte enthalten Seitenäste, die strukturell von der Hauptkette verschieden sind. Pfropfpolymere bestehen wie ein Blockpolymer aus verschiedenartigen Blöcken, doch sind diese nicht endständig verbunden; die Pfropfäste sind vielmehr wie Borsten an der Hauptkette fixiert (Abb. 5.4.4-7). Pfropfpolymere lassen sich aus Lösungen von Homopolymeren in einem Fremdmonomer herstellen. Die gelösten Makromoleküle bilden nach entsprechender Aktivierung Ausgangszentren für die Polymerisation der Pfropfäste. Die Anregung kann rein chemisch oder durch Bestrahlung erfolgen. Verteilung und Länge der Seitenketten sind im allgemeinen statistisch. Pfropfpolymere sind technisch interessant, weil sie gezielte Veränderungen der Werkstoffeigenschaften ermöglichen, die sich auf die Oberfläche des Materials beschränken lassen. So kann z. B. die Benetzbarkeit oder die elektrische Leitfähigkeit gesteigert werden, oder es können durch Bindung ionogener Gruppen an Folien aus zähfesten Materialien Ionenaustauschermembranen hergestellt werden.

5.4.4.6. Vernetzte Polymere. Aufbaureaktionen unter Einbeziehung von polyfunktionellen Grundmolekülen führen bei höheren Umsätzen zu dreidimensional vernetzten Produkten, weil die Seitenketten zusammenwachsen. Da alle Teile durch chemische Bindungen miteinander verankert sind, können vernetzte Stoffe prinzipiell nicht ohne Abbau schmelzen oder in Lösung gehen. Unter gewissen Voraussetzungen nehmen sie Lösungsmittel auf und quellen. Die Vernetzungsdichte kann durch das Verhältnis der Anzahl Grundbausteine, die Vernetzungsstellen bilden, zur Anzahl der insgesamt vorhandenen Grundbausteine ausgedrückt werden.

In der Regel werden vernetzte Polymere in zwei Stufen hergestellt. Nach der ersten wird ein Vorprodukt erhalten, das noch nicht vernetzt und daher noch fließfähig ist. Nach der Formgebung wird die Vernetzung (Härtung) vollzogen.

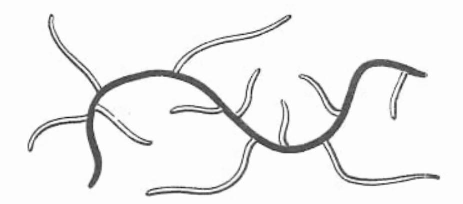


Abb. 5.4.4-7 Pfropfpolymer mit zwei Arten von Blöcken

Kleinere vernetzte Teilchen (Gelkörner) werden in einem Zuge durch vernetzende Aufbaureaktionen, z. B. Copolymerisation mit multi-

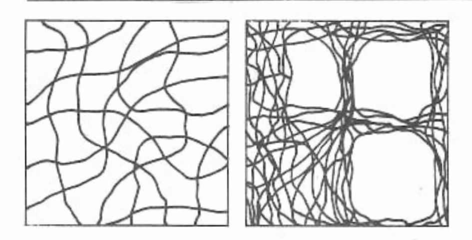


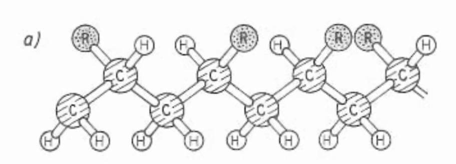
Abb. 5.4.4-8 a) homogen- und b) makroporös vernetzte Gelteilchen

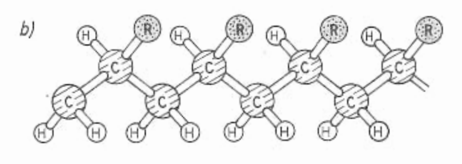
funktionellen Monomeren, hergestellt. In Gegenwart von nicht reagierenden Lösungsmitteln können dabei makroporös vernetzte Strukturen entstehen (Abb. 5.4.4-8), die kaum quellen und dadurch sehr druck- und dimensionsstabil sind, andererseits aber günstige Hohlraumstrukturen zur Aufnahme von Probenmolekülen bei der GPC aufweisen (↑ 5.4.3.2.4.).

5.4.5. Konfiguration der Makromoleküle

Trotz gleicher Verknüpfung der Atome zum Makromolekül, trotz gleicher *Konstitution* (↑ 5.4.4.) können Polymere mit vierwertigen Gerüstatomen in ihren Eigenschaften Unterschiede zeigen, die auf Unterschieden im räumlichen Bau beruhen. Es handelt sich um Effekte, die der Stereoisomerie von niedermolekularen Verbindungen entsprechen (↑ 5.2.2.2.). Bei der Umwandlung von Konfigurationsisomeren müssen chemische Bindungen gesprengt und neue geknüpft werden. Die Aktivierungsenergie ist so hoch, daß die Isomere bei gewöhnlicher Temperatur beständig sind.

5.4.5.1. Taxie. In 1,2-Vinylpolymeren trägt jedes zweite Kohlenstoffatom der Hauptkette einen Substituenten. Denkt man sich das Makromolekül zu einer ebenen Zick-Zack-Kette ausgerichtet, so erkennt man für die Anordnung der Substituenten bezüglich der Ebene, in der die Kohlenstoffatome der Hauptkette liegen, drei Möglichkeiten (Abb. 5.4.5-1): Bei *ataktischer Anordnung* sind sie statistisch verteilt, in einer *isotaktischen Struktur* liegen sie alle auf einer Seite der C—C-Ebene, bei *syndiotaktischer Struktur* dagegen abwechselnd links oder rechts dieser Ebene. Die Taxie wird in der Aufbaureaktion festgelegt. Die stereospezifische Polymerisation, z. B. zu einem isotaktischen Polymer, erfordert Ausrichtung eines jeden Grundmoleküls vor seiner Anlagerung. Das kann durch spezielle Polymerisationskatalysatoren erreicht werden, z. B. durch feste, feindisperse Komplexe aus Aluminiumtrialkyl und Titan(III)-chlorid. Isotaktische und ataktische





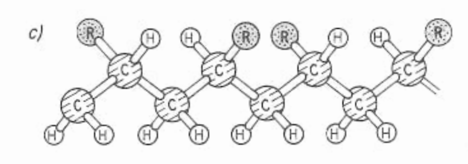


Abb. 5.4.5-1 Stereoregularität bei Vinylpolymeren; a) ataktisch, b) isotaktisch, c) syndiotaktisch

Abb. 5.4.5-3 a) 1,4-cis-Polyisopren (Naturkautschuk), b) 1,4-trans-Polyisopren (Guttapercha)

Tab. 5.4.5-2 Vergleich der Eigenschaften von ataktischen und isotaktischen Polymeren; *a* ataktisch, *i* isotaktisch

Polymer	Erweichungspunkt (Schmelzpunkt) in °C		Dichte in $g\,cm^{-3}$	
Konfiguration	*a*	*i*	*a*	*i*
Polypropylen	75	160	0,85	0,92
Polybuten-1	65	128	0,87	0,91
Polystyren	100	230	1,06	1,08

Polymere unterscheiden sich in wesentlichen Eigenschaften (Tab. 5.4.5-2).

5.4.5.2. cis-trans-Isomerie. Wie in der niedermolekularen Chemie können bei Polymeren mit Doppelbindungen in der Kette Konfigurationsunterschiede durch die Stellung der Substituenten an den doppelt gebundenen Atomen auftreten. So ist Naturkautschuk 1,4-cis-Polyisopren, Guttapercha hingegen 1,4-trans-Polyisopren (Abb. 5.4.5-3). Die cis-Konfiguration kann durch Polymerisation von Isopren mit Triisobutylaluminium bei 50 °C erhalten werden. Das Verfahren wird technisch durchgeführt und liefert ein Produkt, das in seinen Eigenschaften mit Naturkautschuk übereinstimmt. Auch bei Makromolekülen mit alicyclischen Gerüstbausteinen kann cis-trans-Isomerie auftreten.

5.4.6. Konformation der Makromoleküle

Konformationsisomere sind durch niedrige Energieschwellen getrennt, die bei gewöhnlicher Temperatur durch die Energie der Wärmebewegung überwunden werden. Beim Ethan sind 11,7 $kJmol^{-1}$ erforderlich, um eine Methylgruppe durch die Stellung hindurchzudrehen, in der ihre H-Atome die der anderen CH_3-Gruppe gerade verdecken. Bei einer vollen Umdrehung sind drei solche Barrieren zu überwinden, zwischen denen die Energieminima der gestaffelten Konformationen liegen (Abb. 5.4.6-1). Übernimmt man diese Vorstellung für Polymere mit Kohlenstoffgerüst, so hat jede kettenaufbauende C—C-Bindung drei energetisch begünstigte Lagen gegenüber den vorausgehenden. Das führt bei 1000 C-Atomen zu $3^{997} \approx 5 \cdot 10^{475}$ Möglichkeiten. Eine davon ist die gestreckte Kette, die ohne äußeren Zwang folglich kaum realisiert wird.

In der Kohlenstoffkette ist die *trans-Stellung* gegenüber den beiden anderen gestaffelten Stellungen *gauche* + und *gauche* −, je nach Drehsinn, etwas begünstigt (Abb. 5.4.6-2).

5.4.6.1. Konformation in Lösung. In verdünnten Lösungen nehmen Makromoleküle Knäuelgestalt an, die durch den *Fadenendenabstand R* oder den *Trägheitsradius S* charakterisiert werden kann. Für den Fall ungehinderter Rotation (Index f) und Gleichwertigkeit aller drei gestaffelten Positionen kann aus einer geometrischen Betrachtung der Kette mit dem Valenzwinkel 109,5° und der Bindungslänge 0,154 nm für den quadratischen Mittelwert $\langle R^2 \rangle_{0,f}^{0,5} = 0{,}308\sqrt{P}$ (in nm) abgeleitet werden.

Die Knäueldimensionen steigen also mit der Wurzel aus dem Polymerisationsgrad. Beim realen Knäuel tritt ein *Behinderungsparameter* $\sigma = (\langle R^2 \rangle_0 / \langle R^2 \rangle_{0,f})^{0,5}$ mit Werten zwischen 1,3 und 2,7 auf, der die Behinderung der Drehbarkeit berücksichtigt. Der Wert $\langle R^2 \rangle_0$ gilt für die *ungestörten Dimensionen*, die im *Θ-Zustand* vorliegen, in dem der 2. Virialkoeffizient (↑ 5.4.3.4.5.) den Wert $A_2 = 0$ und der Exponent der Viskositätsgleichung (↑ 5.4.3.5.1.) den Wert $a = 0{,}5$ hat. Gute Lösungsmittel weiten die Knäuel auf, was durch den Expansionskoeffizienten $\alpha = (\langle R^2 \rangle / \langle R^2 \rangle_0)^{0,5}$ berücksichtigt wird. So gilt für den quadratischen Kettenabstand insgesamt $\langle R^2 \rangle = \alpha^2 \langle R^2 \rangle_0 = \alpha^2 \sigma^2 \langle R^2 \rangle_{0,f}$. Der Trägheitsradius ist kleiner als der Fadenendenabstand. Für den Θ-Zustand gilt $\langle R^2 \rangle_0 = 6 \langle S^2 \rangle_0$. Die mit den Spitzklammern $\langle \rangle$ ausgedrückte Mittelwertbildung bezieht sich auf die Werte von Knäuel zu Knäuel, auf die zeitliche Änderung durch Konformationsübergänge und auf die Knäuelgestalt.

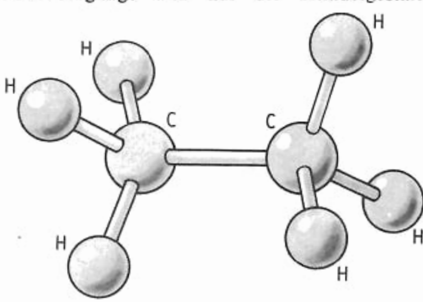


Abb. 5.4.6-1 Ethanmolekül

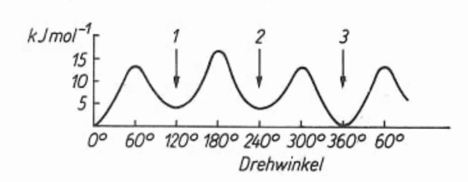


Abb. 5.4.6-2 Konformationsenergie von Butan; *1*, *2* gauche-, *3* trans-Konformation

Makromoleküle in Lösung haben nicht Kugelform, sondern sind *Ellipsoide* mit dem Achsenverhältnis 1 : 1,8 : 3,8. Der Kettenendenabstand kann aus der Winkelabhängigkeit des Streulichts bestimmt werden (↑ 5.4.3.4.5.). Die dort eingeführte Größe ist $d = \langle R^2 \rangle^{0,5}$. Die *Knäueldichte* läßt sich aus Viskositätsmessungen und aus Untersuchungen mit der Ultrazentrifuge bestimmen. Sie beträgt wenige Promille der Dichte des kompakten Polymermaterials. Die Knäuel sind also außerordentlich lockere Gebilde. Mit steigendem Polymerisationsgrad nimmt die Knäueldichte ab. Verzweigte Moleküle haben höhere Werte als lineare. So bietet die Knäueldichte eine Möglichkeit, den *Verzweigungsgrad* (↑ 5.4.4.4.) zu ermitteln.

5.4.6.1.1. *Geordnete Strukturen in Lösung.* Manche Polymere, z. B. Polycarbonat, Cellulosenitrat, neigen dazu, unter bestimmten Bedingungen bei längerem Stehen aus ihren Lösungen auszukristallisieren. Vermutlich erfolgt bereits in der Lösung eine Vorordnung der Moleküle, die der Konformation im Kristall ähnelt (↑ 5.4.6.2.1.).
In einigen Fällen konnten Ordnungszustände in Lösung experimentell nachgewiesen werden. Adenyl-Ribose-Polyphosphat, ein Polynukleotid, bildet z. B. in sauren wäßrigen Lösungen Spiralstrukturen, indem sich zwei Moleküle zu einer doppelgängigen Schraube umeinander winden. Durch Beeinflussung der Kräfte zwischen den Makromolekülen, im vorliegenden Fall durch Neutralisation der Lösung, zerfallen diese *Helixstrukturen* und die Moleküle gehen reversibel in die Knäuelkonformation über. Polyglutaminsäure erfährt eine Helix-Knäuel-Umwandlung bei $pH = 5$, Polyprolin in *n*-Butanol durch Zusatz von 45% Bencylalkohol zum Lösungsmittel.

5.4.6.2. Konformation im lösungsmittelfreien Zustand. Viele Polymere sind glasklar, d. h., sie enthalten keine optischen Inhomogenitäten in der Größenordnung der Lichtwellenlänge. Physikalische Untersuchungen bestätigen, daß sie nicht kristallisiert sind. Ob ein Polymer im festen Zustand kristallin oder amorph ist, hängt von der Molekülstruktur ab.

5.4.6.2.1. *Kristalline Polymere.* Regelmäßig gebaute Ketten können sich parallel lagern und dicht gepackte, geordnete Bereiche bilden. Polyethylenketten und die Methylensequenzen von Polyamiden haben in diesen Kristalliten all-trans-Konformation. Auch syndiotaktische Vinylpolymere nehmen eine solche Zickzackstruktur an, in der sich die Substituenten nicht gegenseitig stören. Bei isotaktischen Vinylpolymeren verhindern die Substituenten die Ausbildung von all-trans-Sequenzen. Hier erfolgt die Kristallisation in gauche-Konformation; die Einzelketten bilden Helices. Von Größe und Beschaffenheit der Substituenten hängt die Steigung dieser Helices ab. Stabil ist die Struktur, bei der die Summe von intra- und intermolekularer Energie ein Minimum hat. In der graphischen Darstellung der Konformationsenergie wird das Potentialfeld in den Koordinaten der Winkel φ' und φ'' für Drehungen um sukzessive C—C-Bindungen gezeichnet (Abb. 5.4.6-3). Aus der Lage des Minimums ergibt sich die stabile Konformation, z. B. für isotaktisches Polypropylen die 3_1-Helix.

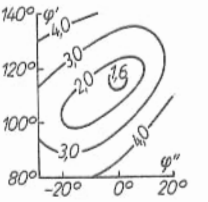


Abb. 5.4.6-3 Konformationsenergie als Funktion der Drehung um die Winkel φ' und φ'' für isotaktisches Polypropylen

5.4.6.2.2. *Amorphe Polymere.* Unregelmäßig gebaute Makromoleküle können nicht kristallisieren. Das gilt für die meisten Copolymere, für ataktische Polymere und stark verzweigte Moleküle. Sie erstarren glasartig in Knäuelkonformation mit den ungestörten Dimensionen. Diese Aussage der Theorie ist seit 1973 durch Kleinwinkelstreuung mit langsamen Neutronen von mehreren Arbeitsgruppen experimentell bestätigt worden.

Tab. 5.4.7-1 Aminosäuren, die am Aufbau der meisten Proteine teilnehmen

Glycin	Gly
Alanin	Ala
Valin	Val
Leucin	Leu
Isoleucin	Ile
Serin	Ser
Threonin	Thr
Cystein	Cys
Methionin	Met
Asparaginsäure	Asp
Glutaminsäure	Glu
Glutamin	Gln
Asparagin	Asn
Arginin	Arg
Lysin	Lys
Phenylalanin	Phe
Tyrosin	Tyr
Tryptophan	Try
Histidin	His
Prolin	Pro

5.4.7. Strukturaspekte von Biopolymeren

5.4.7.1. Individualität und Strukturspezifität der Biopolymeren. Makromolekularen Naturprodukten, wie Zellulose, Stärke, Kautschuk, Nukleinsäuren, Enzymen und Strukturproteinen, kommt im biologischen Geschehen zentrale Bedeutung zu. Die Makromoleküle haben eine *homo*- oder eine *copolymere Struktur*, je nachdem ob sie aus gleichen oder verschiedenen monomeren Grundmolekülen aufgebaut sind; bei copolymerer Struktur können die Monomeren in unregelmäßigen oder in regelmäßigen Sequenzen aufeinander folgen. Damit ist ihre Struktur schon so komplex, daß jedem Makromolekül eine Individualität zukommt, die für die biologischen Strukturen lebender Organismen wichtig ist. Die biologisch wichtigsten Systeme dieser Art sind *Proteine* und

Enzyme, die aus rund 20 verschiedenen Aminosäuren in spezifischer Anordnung bestehen (Tab. 5.4.7-1), sowie *Desoxyribo-* und *Ribonukleinsäuren*, die aus je vier verschiedenen Mononukleotiden (↑ 5.4.7.3.) ebenfalls in spezifischen Sequenzen aufgebaut sind. Die *Primärstruktur* der Makromoleküle wird durch die Anordnung der Monomeren bestimmt; das Insulinmolekül z. B. besteht aus einer *A*- und einer *B*-Kette mit 21 bzw. 30 Aminosäuren in spezifischer Reihenfolge (Abb. 5.4.7-2).

Unterschiede in Einzelheiten der Struktur und der Form der Proteine funktionell und strukturell gleichen Typs sind die Ursache der *Mikroheterogenität* der Eiweiße zwischen verschiedenen biologischen Arten und darüber hinaus zwischen verschiedenen Individuen der gleichen Art. Die Sequenzen der Primärstruktur bestimmen die Eigenschaften des jeweiligen Biopolymers. Die Folge der Monomeren enthält die für das Leben des biologischen Systems notwendige *Information* und entspricht der Verknüpfung von Buchstaben des Alphabets zu Wörtern, Sätzen und Texten. Durch innere Drehungen von Teilen der Makromoleküle gegeneinander ergeben sich *Überstrukturen*. *Sekundäre Strukturen* entstehen durch Streckung, Knäuelung und Faltung der Ketten, ohne daß dabei die Primärstruktur gestört wird. Durch räumliche Ausrichtung und Orientierung von Molekülgruppen bestimmt die Überstruktur zugleich die Art der Wechselwirkung mit Nachbarmolekülen und fixiert sie räumlich. *Zwischenmolekulare Bindungen* ergeben eine Assoziation der linearen Polymere zu hoch orientierten Aggregaten. Die Anisotropie und der hohe Ordnungsgrad biologischer Strukturen sind Voraussetzungen für die Eigenschaften und die Funktionsweisen lebender Systeme.

Sowohl Eiweiße als auch Nukleinsäuren stellen *Polyelektrolyte* dar. Zwischen ihren einzelnen Gliedern wirken elektrostatische Kräfte, deren Potential durch den Ionisationsgrad bestimmt wird. Ein geringer Ionisationsgrad begünstigt eine Knäuelung des Moleküls, während ein hoher Ionisationsgrad wegen der dann auftretenden Abstoßungskräfte eine Streckung des Moleküls zu einer planaren Zickzackform oder zu einer räumlichen Spiralform bewirkt.

5.4.7.2. Konformation der Proteine. Die Sekundärstrukturen rotationsisomerer Makromoleküle werden über zwischenmolekulare Wechselwirkungen polarer Gruppen stabilisiert, die *Polypeptidmoleküle* der Proteine z. B. über Wasserstoffbrücken zwischen Gruppen N—H...O=C=. Die größte Stabilität wird bei linearer Anordnung der Atome N—H ...O=C< erreicht und damit zugleich die Anzahl der möglichen stabilen Konstellationen stark eingeschränkt. Es treten nur zwei Molekültypen mit *Schraubenstruktur*, mit schraubenartiger Anordnung der Aminosäuren auf, die γ-, und die *α-Helix*. Die γ-Helix hat einen großen Hohlraum entlang der Achse, und ihre Struktur ist wenig stabil; die α-Helix dagegen ist eine stereochemisch und energetisch günstigere Konformation, in ihr hat die Polypeptidkette die dichteste Anordnung um die Schraubenachse bei minimalem Durchmesser einer Helixwindung, so daß die zwischenmolekularen Kräfte groß sind. Die Ganghöhe der Helix beträgt rund 54 nm, auf eine Helixwindung entfallen 3,6 Aminosäurereste. Jede Gruppe N—H ...O=C< ist über Wasserstoffbrücken mit der ihr drittnächst benachbarten Säureamidgruppe in beiden Richtungen verbunden. Aus theoretischen Gründen ist für eine α-Helix, deren Primärstruktur nur *L*-Aminosäuren enthält, die Rechtsschraube gegenüber der Linksschraube stereochemisch und energetisch begünstigt. Aus der Beugung von Röntgenstrahlen und der Doppelbrechung im Infrarotbereich ergibt sich, daß viele synthetische Polypeptide und natürliche Proteine zu einem großen Teil α-Helixstruktur haben.

Außer der Helixanordnung kann auch eine gegenseitige Fixierung benachbarter Linearpolymeren zu einer periodischen Struktur führen, eine parallele flächenartige Anordnung von Polypeptidketten z. B. zu *Faltblattstrukturen*. Eine Aggregation mehrerer Ketten zu stärkeren Fasern oder Rückfaltungen und

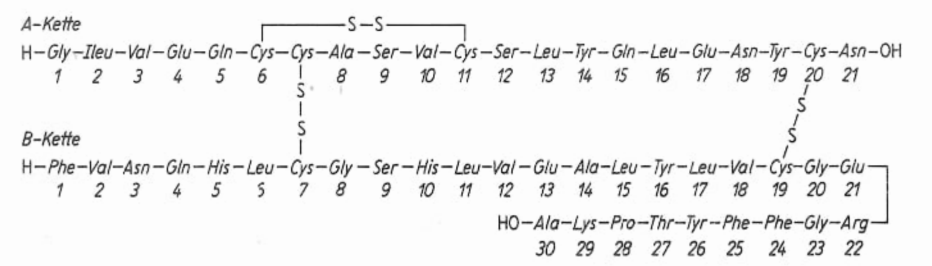


Abb. 5.4.7-2 Primärstruktur des Rind-Insulins (nach PERUTZ)

Knäuelungen der vollständig oder teilweise spiralisierten einzelnen Helixmoleküle zu kugelförmigen Strukturen verkörpert die *Tertiärstruktur* der globulären Proteine; beim Kollagen z. B. sind 3 spiralige Polypeptidketten umeinander verdrillt. Die Faltungen schließen Wechselwirkungen der Seiten- und Hauptketten untereinander sowie die Wirkung verschiedener Nebenvalenzen ein. Native globuläre Proteine (Enzyme) sind oft aus mehreren gleichen oder verschiedenen Polypeptiden aufgebaut. Wechselwirkungen zwischen solchen Untereinheiten ergeben eine *quaternäre Struktur*. Molekulare Wechselwirkungen können über strukturverändernde Einflüsse gestört werden. Bei einer *Denaturierung* werden die stabilisierenden Wasserstoffbindungen gelöst. Sie ist eine spezifische Eigenschaft biologisch aktiver Makromoleküle und führt stufenweise durch Strukturveränderungen zum Verlust der biologischen Aktivität der nativen Struktur. Die ersten Stufen haben dabei reversiblen Charakter; eine vollständige Denaturierung stellt aber eine *irreversible Strukturveränderung* dar.

5.4.7.3. Sekundäre Strukturen der Nukleinsäuremoleküle. Ebenso wie Polypeptide zeigen auch die *Polynukleotidketten* der Nukleinsäuremoleküle in ihrer Primärstruktur Polarität und Richtungssinn. Die *Mononukleotide* als monomere Elemente stellen Verbindungen einer *Purin-* oder einer *Pyrimidinbase* mit einem *Zuckermolekül* mit 5-C-Atomen dar, dessen 5′-C-Atom mit Phosphorsäure verestert ist. Am Aufbau der Desoxyribonukleinsäure- oder *DNS-Mononukleotide* nehmen neben Phosphorsäure die Purinbasen *Adenin* (A) oder *Guanin* (G) bzw. die Pyrimidinbasen *Thymin* (T) oder *Cytosin* (C) und 2′-Desoxyribose teil. In Ribonukleinsäure- oder *RNS-Mononukleotiden* tritt anstelle des Thymin *Uracil* (U) auf und 2′-Deoxyribose ist durch Ribose ersetzt. Benachbarte monomere Grundmoleküle sind über Phosphordiesterbindungen zwischen Hydroxylgruppen in 3′- und 5′-Stellung der Zuckerkomponenten miteinander verknüpft, so daß ein Ende des Polynukleotids eine freie Hydroxylgruppe am 3′-C-Atom des Zuckers, das andere Ende am 5′-C-Atom des Zuckers eine Phosphorsäure trägt. Entsprechend wird die Kettenpolarität bei Proteinen durch Amino- und Carboxylendgruppen bedingt.

DNS-Moleküle sind aus langen, unverzweigten Fadenmolekülen aufgebaut. Ihr Molekulargewicht schwankt in Abhängigkeit vom Darstellungsverfahren, da ebenso wie für RNS-Moleküle nicht immer zwischen Aggregaten und Einzelmolekülen unterschieden werden kann. Aus Mikroorganismen wurden DNS-Moleküle mit einem Molekulargewicht von 10^8 bis 10^9 isoliert. Der Gehalt $(A + G)$ und damit die Anzahl der Purinbasen pro DNS-Molekül ist stets der Anzahl $(C + T)$ der Pyrimidinbasen gleich; dabei treten A und T bzw. G und C gleich häufig auf, und die Anzahl der Basen, die am 6-C-Atom des Basenrings eine NH_2-Gruppe tragen, ist der Anzahl Basen gleich, die am 6-C-Atom eine CO-Gruppe tragen. Die molekulare Zusammensetzung der DNS verfügt nur über einen Freiheitsgrad $(A + T)/(G + C)$. Aus den Beugungsdiagrammen von Röntgenstrahlen ergibt sich eine Spiralstruktur der DNS-Moleküle. Nach den Daten für Durchmesser und Ganghöhe der Spirale sowie für Länge und Masse des Moleküls liegt eine Doppelspirale vor. Nach WATSON und CRICK bilden zwei Polynukleotidketten mit gegenläufiger Polarität eine Doppelspirale, in der sie aufsteigend entgegen dem Uhrzeigersinn um eine gemeinsame Molekülachse geschlungen sind. Dabei sind die beiden Enden der Doppelspirale im Unterschied zu den Enden der einzelnen Ketten strukturell gleichberechtigt. Die vorliegende Struktur der Doppelspirale wird über Wasserstoffbrücken zwischen NH_2-Gruppen und CO-Gruppen einander gegenüberliegender Basen stabilisiert; dabei können aus sterischen Gründen nur A—T- und G—C-Paarungen zwischen den beiden Einzelketten auftreten. Die sich paarenden Basen liegen in der Helix nicht diametral einander gegenüber, so daß die Wasserstoffbrücken sich seitlich von der Zentralachse befinden. Zwischen A—T-Paaren werden zwei, zwischen G—C-Paaren drei Wasserstoffbrücken ausgebildet. Die beiden Einzelstränge in dem α-Doppelhelixmolekül der DNS haben bezüglich ihrer Basensequenz komplementären Charakter. Die Sequenz der einen Kette determiniert über spezifische Basenpaarungen die Sequenz der anderen. Die Basen liegen im Hohlraum der Doppelhelix planar mit einem Ringebenenabstand von 0,34 nm übereinander und bilden mit der Molekülachse einen Winkel von 90°, während die infolge einer Dissoziation der Phosphatgruppen negativ geladene Zucker-Phosphorsäure-Kette die Grenzfläche des Moleküls zum umgebenden Medium bildet. Hydrophobe Attraktionskräfte zwischen übereinanderliegenden Basen in Richtung der Molekülachse tragen zur Strukturstabilisierung bei. Beide Einzelstränge sind weniger als eine halbe Windung gegeneinander versetzt. Die Parameter der Sekundärstruktur der DNS-Doppelhelix ändern sich in Abhängigkeit vom Wassergehalt, von der Ionenart und -konzentration des umgebenden Mediums. Die *B-Form* der DNS im lebenden System hat eine Identitätsperiode von 0,346 nm; 10 Basen bilden in ihr bei senkrechter Anordnung um die Molekülachse eine Helixwindung. Bei weniger als 70% relativer Feuchtigkeit kristallisiert das Na-Salz der DNS in der *A-Form*, bei der 11 Nukleotide eine Helixwindung der Ganghöhe 0,280 nm bilden und etwa um 70°

gegen die Molekülachse geneigt angeordnet sind. Die kristalline *C-Form* der DNS erhält man bei weniger als 44% relativer Feuchtigkeit in Form des Lithiumsalzes. Hier bilden 9,3 Nukleotide eine Helixwindung der Ganghöhe 0,309 nm.

Weder in Gelen noch in verdünnten Lösungen zeigen die Doppelhelixmoleküle der DNS im Unterschied zu Proteinen Tendenzen zur Assoziation. Im lebenden System sind DNS-Moleküle über salzartige Bindungen an basische Proteine gebunden, die in Form eines Strangs als dritte koaxiale Komponente den Zwischenraum in der Doppelhelix zwischen beiden Polynukleotidsträngen ausfüllen und die Struktur stabilisieren. Durch Proteinbrücken zwischen verschiedenen DNS-Molekülen wird deren räumliche Anordnung fixiert, so daß die Nukleoproteide höhere Ordnungszustände darstellen als die freien Doppelhelixmoleküle der DNS. Zugleich können sich über die Proteinkomponente Tertiär- und Quartärstrukturen ausbilden, da die DNS-Moleküle im lebenden System in Strukturen kleinerer Dimensionen angeordnet sind, als sich aus dem Watson-Crick-Modell für eine starre Doppelhelix ergeben würde. In Lösung entspricht die Konformation des DNS-Moleküls einem *steifen Knäuel* von der Form eines wurmähnlichen Kettenmoleküls.

Die RNS bildet in vivo vorwiegend einsträngige Polymere von unterschiedlicher Kettenlänge mit einem Molekulargewicht von 10^4 bis 10^6. Über Rückfaltungen der polymeren Kette und Basenpaarungen können Teilbereiche des Makromoleküls DNS-ähnliche Ordnungszustände annehmen. Die Röntgendiagramme zeigen Ähnlichkeit mit der *A*-Form der DNS. In Lösung verhält sich das RNS-Molekül entsprechend seiner Einstrangnatur wie ein statistisches Knäuel. Für RNS-Nukleoproteide dagegen trifft das Stäbchenmodell zu. Möglicherweise legen die assoziierten Proteine spezielle Strukturen der RNS fest. Auch die Einstrang-DNS einiger primitiver Bakteriophagen zeigt ein ähnliches Verhalten. Irreversible Strukturveränderungen bei der DNS durch Unter- oder Überschreiten kritischer *p*H-Werte stellen wie bei den Proteinen eine Denaturierung dar, die zur Bildung eines Helixknäuels führt und bei der beide Einzelstränge getrennt werden. Auch Veränderungen der Temperatur, der Ionenstärke und der Dielektrizitätskonstante der Lösung wirken denaturierend. Durch Hitze denaturierte DNS gleicher oder eng verwandter Mikroorganismen können unter besonderen Versuchsbedingungen renaturiert werden und ihre biologische Aktivität wieder erlangen.

Die gleichen Denaturierungsbedingungen führen infolge intramolekularer Wasserstoffbrückenbildung zwischen komplementären oder nahezu komplementären Sequenzbereichen bei der RNS zu einer sehr starken Verknäuelung des einstrangigen Polyelektrolyten.

5.4.7.4. Helixstrukturen synthetischer Polynukleotide. Eine den Biopolymeren äquivalente Sekundärstruktur läßt sich an synthetisch hergestellten Polynukleotiden nachweisen. Das RNS-artige homopolymere Makromolekül *Polyadenylsäure* bildet über Wasserstoffbrücken aus zwei gleichläufigen Polyadenylsäuresträngen eine symmetrische Doppelhelix; ebenso das Homopolymer *Polycytidylsäure.* Die Sekundärstruktur der *Polyinosinsäure* repräsentiert dagegen eine Dreifachhelix aus drei gleichläufigen Polyinosinsäureketten. Sekundärstrukturen des Watson-Crick-Typus bilden sich auch zwischen verschiedenen Homopolymeren. *Polyuridylsäure* und *Polyadenylsäure* bilden über *A—U*-Paarungen einen 1:1-Komplex $Poly(A + U)$, dessen Röntgendiagramm dem der DNS sehr ähnlich ist. Der entsprechende 2:1-Komplex stellt eine Dreifachhelix dar. Der Komplex $Poly(A + I)$ aus Polyadenylsäure und Polyinosinsäure entspricht in seinen Strukturparametern der *B*-Form der DNS. Das Makromolekül $Poly(A + G + C + U)$ mit statistischer Basensequenz hat RNS-Eigenschaften. Polydesoxyribonukleotide bilden den genannten Polyribonukleotiden entsprechende Komplexe.

Nukleoproteide stellen Bindeglieder zwischen strukturierter lebender und unbelebter Materie dar. Sie verkörpern die alleinige molekulare Komponente der einfachsten Viren, die zwar noch als unbelebte Materie angesehen werden können, aber in geeigneter Umwelt Eigenschaften lebender Systeme zeigen.

6. Zustandsformen der Stoffe

6.1. Makroskopische Teilchensysteme

6.1.1. Makroskopische Erscheinungsformen der Materie

Die Bausteine der Materie, Elementarteilchen und Atomkerne, Atome und Moleküle, sind wegen ihrer geringen räumlichen Ausdehnung nicht direkt sinnlich wahrnehmbar. Nur in Sonderfällen ist es gelungen, durch eine mehr als millionenfache Vergrößerung mit dem Feldelektronenmikroskop die räumlichen Umrisse größerer Moleküle sichtbar zu machen (Tafel 30). Die meisten dieser Bausteine werden auf indirektem Wege durch ihre Wechselwirkung mit anderen Teilchen, z. B. durch die Spuren beobachtet, die sie bei ihrer Bewegung hinterlassen. Im allgemeinen hat man es jedoch in der Natur, in der Technik und im Laboratorium nicht mit einzelnen Materieteilchen, sondern mit einer meist direkt wahrnehmbaren Anhäufung einer großen Anzahl von Atomen, Ionen oder Molekülen, d. h. mit einem makroskopischen Teilchensystem, zu tun; z. B. beträgt die Zahl der Einzelteilchen in 1 cm^3 bei 20 °C und 10^5 Pa Druck bei einem Gas etwa 10^{19}, bei einer Flüssigkeit etwa 10^{22} und bei einem festen Stoff etwa 10^{23}.

6.1.1.1. Aggregatzustände und Phasen. Bei den meisten Atom- und Molekülanhäufungen unterscheidet man je nach der Stärke des Zusammenhalts der Teilchen *feste Stoffe* (↑ 6.4.), *Flüssigkeiten* (↑ 6.3.) und *Gase* (↑ 6.2.). Welcher dieser *Aggregatzustände* jeweils vorliegt, hängt molekulartheoretisch gesehen außer von der Stärke der Anziehungskräfte, die die Teilchen aufeinander ausüben, von der Intensität ihrer Wärmebewegung ab, die den Anziehungskräften entgegenwirkt. Im Grenzzustand des *idealen Festkörpers* (↑ 8.1.1.) sind allein die Anziehungskräfte, im *idealen Gaszustand* (↑ 6.2.2.1.) allein die Wärmebewegungen wirksam. Jede Anhäufung von Teilchen, die makroskopisch homogen ist, d. h. räumlich konstante Beschaffenheit aufweist, bezeichnet man als eine *Phase* oder ein *homogenes System.* Eine *reine Phase* oder ein *reiner Stoff* besteht aus lauter gleichartigen Teilchen, die auch teilweise dissoziiert oder assoziiert sein können, eine *Mischphase* oder *Mischung* dagegen aus zwei oder mehreren Stoffen, z. B. eine wäßrige Zuckerlösung aus Wassermolekülen und Zuckermolekülen. Stoffliche Systeme, z. B. die Erdatmosphäre, bei denen sich Größen, wie Druck, Dichte, Konzentrationen u. a., stetig im Raum ändern, stellen *quasihomogene Phasen* oder *kontinuierliche Systeme* dar.

In der materiellen Welt hat man es vielfach mit Systemen zu tun, die sich aus zwei oder mehreren Phasen zusammensetzen. Solche Systeme nennt man *Mehrphasensysteme* oder *heterogene Systeme*, z. B. können bei bestimmten Temperaturen und Drücken flüssiges Wasser und festes Eis nebeneinander vorhanden sein. Die Berührungsflächen zwischen den einzelnen Phasen, an denen sich die Eigenschaften unstetig ändern, werden *Phasengrenzflächen* (↑ 6.6.1.) genannt.

6.1.1.2. Methoden zur theoretischen Ermittlung des makroskopischen Verhaltens der Materie. Für die theoretische Erforschung und Beschreibung von makroskopischen Vielteilchensystemen stehen in der Hauptsache zwei Methoden zur Verfügung, die thermodynamische und die statistische Methode.

Die *thermodynamische Methode* geht von ganz wenigen grundlegenden Axiomen aus, von Erfahrungssätzen, deren Gültigkeit nur experimentell zu bestätigen ist. Diese Sätze mit sehr allgemeinen Aussagen über das Verhalten der Materie werden als *Hauptsätze der Thermodynamik* bezeichnet. Mit ihrer Hilfe lassen sich einige charakteristische Größen, wie innere Energie oder Entropie, festlegen, die für jedes Teilchensystem in bestimmter Weise von solchen Bedingungen, wie Druck, Temperatur und stofflicher Zusammensetzung, abhängen. Daraus ergeben sich dann konkrete Zusammenhänge zwischen den verschiedensten beobachtbaren makroskopischen Eigenschaften eines Systems. Die Gesamtheit dieser Zusammenhänge bildet ein in sich abgeschlossenes Formelsystem, mit dessen Hilfe das makroskopische Verhalten der Materie vollständig beschrieben werden kann; sind z. B. einige der Eigenschaften eines Teilchensystems aus experimentellen Beobachtungen bekannt, so lassen sich andere aufgrund der thermodynamischen Beziehungen berechnen. Da die *phänomenologische Thermodynamik* nur Zusammenhänge zwischen den makroskopischen Eigenschaften eines Systems liefert, ist es grundsätzlich unmöglich, allein mit ihrer Hilfe irgendwelche Aufschlüsse über die Struktur der Materie und das Verhalten der korpuskularen Bausteine zu erhalten. Die thermodynamischen Zusammenhänge würden auch gelten, wenn die Materie kontinuierlich aufgebaut wäre.

Die *statistische Methode* (↑ 2.4.) geht von diskreten Teilchen aus, wobei das Verhalten einer großen Teilchenzahl untersucht wird. Da sich die Teilchen bezüglich ihrer Geschwindigkeit oder Energie in der Regel keineswegs alle im gleichen Zustand befinden, muß zur Berechnung der makroskopischen Eigenschaften des Gesamtsystems untersucht werden, wie sich die Teilchen auf die verschiedenen möglichen Zustände verteilen. Die Ermittlung dieser *Verteilungsfunktion* (↑ 2.4.4.2.) und ihre Verknüpfung mit den thermodynamischen Größen ist eine Hauptaufgabe der *statistischen Thermodynamik*. Diese Methode liefert streng genommen nur *Wahrscheinlichkeitsaussagen*, die jedoch bei sehr großen Teilchenzahlen schließlich den Charakter von exakten Gesetzen annehmen.

Bei einer speziellen Variante der statistischen Betrachtungsweise, bei der *kinetischen Theorie der Materie* (↑ 6.2.4.), unternimmt man es schließlich, charakteristische mechanische Zustandsgrößen, z. B. Lagekoordinaten, Geschwindigkeit oder Energie, von einzelnen Teilchen zu fixieren und ihre Veränderungen infolge der Bewegungen der Teilchen, der Einwirkung von Nachbarteilchen, von äußeren Kräften u. a. zu verfolgen. Betrachtet man gleichzeitig eine große Anzahl von Teilchen, so gelangt man auch auf diesem Wege wieder zu Aussagen über die relative Häufigkeit der genannten Zustände, d. h. zu einer Verteilungsfunktion.

6.1.2. Phänomenologische Thermodynamik

6.1.2.1. Zustandsgrößen und Zustandsfunktionen. Der *innere Zustand* eines stofflichen Systems, das durch physikalische oder auch nur gedachte Wände gegen seine Umgebung abgegrenzt ist, kann durch meßbare makroskopische Eigenschaften charakterisiert werden. Man unterscheidet extensive und intensive Eigenschaften. *Extensive Eigenschaften*, z. B. Masse, Volumen, eines heterogenen Systems setzen sich additiv aus den entsprechenden extensiven Eigenschaften der einzelnen Phasen zusammen (↑ 6.1.2.1.5.). *Intensive Eigenschaften* behalten ihren Wert, addieren sich also nicht, wenn man gleiche Systeme miteinander vereinigt. Solche Eigenschaften sind z. B. Temperatur, Druck oder Dichte.

Stoffcharakteristische Eigenschaften von 1 g, der Einheit der Masse, bzw. von 1 mol = $6{,}0225 \cdot 10^{23}$ Teilchen, der Einheit der Teilchenmenge, eines Stoffs werden durch *spezifische* bzw. *molare Größen* ausgedrückt. So ist z. B. das Volumen eines Gramms eines reinen homogenen Stoffs das spezifische Volumen und das Volumen eines Mols das molare Volumen oder Molvolumen.

Die Erfahrung zeigt, daß sich der makroskopische Zustand eines *abgeschlossenen Systems*, das mit seiner Umgebung weder Energie noch Materie austauschen kann, nach kürzerer oder längerer Zeit nicht mehr verändert. Das System befindet sich dann in seinem *thermodynamischen Gleichgewichtszustand*.

6.1.2.1.1. *Zustandsgrößen.* Die zur Beschreibung des inneren Zustands eines Systems verwendeten physikalischen Größen nennt man Zustandsgrößen. Bei einem homogenen Stoff genügt häufig die Angabe des Volumens v, des Drucks p und der Temperatur T. Bei Mischphasen aus den k Komponenten $X_1, X_2, \ldots, X_k$ sind darüber hinaus noch Angaben über die stoffliche Zusammensetzung erforderlich. Dazu dienen meist die folgenden *Zusammensetzungsvariablen*:

$n_k = \bar{m}_k/M_k$ Teilchenmenge der Komponente X_k, Einheit: 1 mol;

$x_k = n_k/\sum n_l$ Molenbruch der Komponente X_k;

$w_k = \bar{m}_k/\sum \bar{m}_l$ Massenbruch der Komponente X_k;

$m_k = n_k/\bar{m}_j$ Molalität der Komponente X_k, Einheit: 1 mol kg^{-1};

Tab. 6.1.2-1 Umrechnung von Zusammensetzungsvariablen

n_k	$= x_k \Sigma n_l$	$= w_k \dfrac{\Sigma \bar{m}_l}{M_k}$	$= m_k n_j M_j$	$= c_k v$
$n_k \dfrac{1}{\Sigma n_l}$	$= x_k$	$= w_k \dfrac{1}{M_k \Sigma \dfrac{w_l}{M_l}}$	$= m_k \dfrac{1}{\dfrac{1}{M_j} + \Sigma m_{l \neq j}}$	$= c_k \dfrac{M_j}{\varrho + \Sigma c_{l \neq j}(M_j - M_{l \neq j})}$
$n_k \dfrac{M_k}{\Sigma n_l M_l}$	$= x_k \dfrac{M_k}{\Sigma x_l M_l}$	$= w_k$	$= m_k \dfrac{M_k}{1 + \Sigma m_{l \neq j} M_{l \neq j}}$	$= c_k \dfrac{M_k}{\varrho}$
$n_k \dfrac{1}{n_j M_j}$	$= x_k \dfrac{1}{x_j M_j}$	$= w_k \dfrac{1}{w_j M_k}$	$= m_k$	$= c_k \dfrac{1}{\varrho - \Sigma c_{l \neq j} M_{l \neq j}}$
$n_k \dfrac{1}{v}$	$= x_k \dfrac{\varrho}{\Sigma x_l M_l}$	$= w_k \dfrac{\varrho}{M_k}$	$= m_k \dfrac{\varrho}{1 + \Sigma m_{l \neq j} M_{l \neq j}}$	$= c_k$

$c_k = n_k/v$ Konzentration der Komponente X_k, Einheit: 1 mol l^{-1};

Dabei bedeuten $\bar{m}$ Masse, M Molmasse, $\bar{m}_j$ Masse einer Vorzugskomponente X_j (meist ein Lösungsmittel), v Volumen der Mischphase, $\varrho = \bar{m}/v$ Dichte der Mischphase.

Tab. 6.1.2-1 gibt die Umrechnung der Zusammensetzungsvariablen an. In manchen Fällen ist die Angabe weiterer Zustandsgrößen erforderlich, z. B. die Oberfläche eines flüssigen oder festen Stoffs oder das elektrische Potential einer Phase.

6.1.2.1.2. *Zustandsänderungen.* Ein Vorgang wird in der Thermodynamik *reversibel* oder umkehrbar genannt, wenn alle Veränderungen, die während seines Ablaufs in dem betrachteten System und seiner Umgebung auftreten, dadurch wieder vollständig zum Verschwinden gebracht werden können, daß man den Vorgang genau auf demselben Weg in der umgekehrten Richtung durchführt. Bei einem *irreversiblen* oder nicht umkehrbaren Vorgang kann der Ausgangszustand des Systems hingegen nicht völlig wieder hergestellt werden, ohne daß in der Umgebung Veränderungen zurückbleiben. Reversibel kann eine Zustandsänderung nur dann sein, wenn sich das System während des Vorgangs ständig im Gleichgewicht befindet, der Vorgang muß dann unendlich langsam verlaufen, da ein schnellerer Ablauf eine endliche Ausrückung des Systems aus dem Gleichgewichtszustand erfordert. In der Natur kommen ausschließlich irreversible Prozesse vor. Im Laboratorium läßt sich jedoch für viele Vorgänge ein wenigstens angenähert reversibler Verlauf erreichen.

Zustandsänderungen bei konstanter Temperatur ($dT = 0$) bezeichnet man als *isotherm*, solche bei konstantem Druck ($dp = 0$) als *isobar* und solche bei konstantem Volumen ($dv = 0$) als *isochor*.

6.1.2.1.3. *Zustandsfunktionen.* Jede Eigenschaft Z eines im Gleichgewicht befindlichen Systems kann im allgemeinen als eindeutige Funktion von mehreren anderen Eigenschaften $x, y, z, \ldots$ dargestellt werden: Eine solche Funktion $Z = f(x, y, z, \ldots)$ nennt man *Zustandsfunktion*, die Eigenschaften $x, y, z, \ldots$ die *Zustandsvariablen*. Da Z nur von dem jeweiligen Zustand des Systems, nicht aber davon abhängt, auf welchem Wege das System dorthin gelangt, kann die Änderung $\Delta Z = Z_2 - Z_1 = f(x_2, y_2, z_2) - f(x_1, y_1, z_1) = f(x_2, y_2, z_2) - f(x_2, y_2, z_1) + f(x_2, y_2, z_1) - f(x_2, y_1, z_1) + f(x_2, y_1, z_1) - f(x_1, y_1, z_1)$ auf Änderungen jeweils einer Variablen zurückgeführt werden. Eine infinitesimale Änderung dZ ist somit ein vollständiges Differential $dZ = (\partial Z/\partial x)_{y,z}\,dx + (\partial Z/\partial y)_{x,z}\,dy + (\partial Z/\partial z)_{x,y}\,dz$, das sich additiv aus Anteilen zusammensetzt, die die Änderung von Z bei Änderung nur jeweils einer der Zustandsvariablen $x, y, z, \ldots$ um $dx, dy, dz, \ldots$ angeben.

Die Aufgabe der Thermodynamik besteht darin, bestimmte für das thermische und energetische Verhalten von stofflichen Systemen charakteristische Zustandsfunktionen aufzufinden, zu konkreten Formen dieser Funktionen zu gelangen und daraus vielfältige weitere Aussagen über das makroskopische Verhalten bestimmter Systeme rechnerisch abzuleiten.

6.1.2.1.4. *Thermische Zustandsgleichung.* Darunter versteht man die Zustandsfunktion, die das Volumen v eines Systems in Abhängigkeit von der Temperatur T, dem Druck p und den Teilchenmengen n angibt. Bei einem reinen homogenen Stoff hat sie die allgemeine Form $v = n \cdot f(p, T)$ bzw. $v/n = V = f(p, T)$, wobei V das Molvolumen bedeutet. Damit ergibt sich $v = (\partial v/\partial T)_{p,n}\,dT + (\partial v/\partial p)_{T,n}\,dp + V\,dn$. Den durch das Volumen dividierten Temperaturkoeffizienten des Volumens $1/v(\partial v/\partial T)_{p,n} = \gamma$ bezeichnet man als *thermischen Ausdehnungskoeffizienten*, den durch das Volumen dividierten negativen Druckkoeffizienten des Volumens $-1/v(\partial v/\partial p)_{T,n} = \varkappa$ als *Kompressibilitätskoeffizienten.* Zwischen dem Molvolumen, dem Ausdehnungskoeffizienten und dem Kompressibilitätskoeffizienten besteht der Zusammenhang $dv = \gamma v\,dT - \varkappa v\,dp + V\,dn$.

6.1.2.1.5. *Partielle molare Größen und Gibbs-Duhemsche Gleichung.* Nur bei „idealen" gasförmigen oder flüssigen Mischphasen (↑ 6.3.3.) setzen sich manche extensiven Eigenschaften, z. B. das Volumen, additiv aus den entsprechenden molaren Eigenschaften $Z_1 \ldots Z_k$ der reinen Bestandteile X_1 bis X_k zusammen. In diesen Fällen gilt für konstanten Druck und konstante Temperatur $Z = n_1 Z_1 + n_2 Z_2 + \ldots + n_k Z_k = \sum_{i=1}^{k} n_i Z_i$. In der Regel verhalten sich die extensiven Eigenschaften jedoch keineswegs additiv, weil die Wechselwirkungskräfte zwischen den Teilchen sowie deren Strukturen und Ordnungszustände in einem Gemisch und in den reinen Stoffen meist verschieden sind. Löst man z. B. $MgSO_4$ in 1000 g H_2O auf, so ist das Volumen der Lösung bei kleinen Konzentrationen des Salzes kleiner als das Volumen des reinen Was-

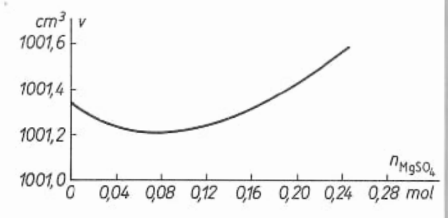


Abb. 6.1.2-2 Volumen v wäßriger Magnesiumsulfatlösungen in Abhängigkeit von der Konzentration bei 18 °C

sers, geht dann durch ein Minimum und nimmt erst bei höheren Konzentrationen stetig zu (Abb. 6.1.2-2). Um auch dann noch eine einfache Additivität $Z = n_1\bar{Z}_1 + n_2\bar{Z}_2 + \ldots + n_k\bar{Z}_k$ zu erhalten, führt man an Stelle der molaren Größen Z der reinen Stoffe die *partiellen molaren Zustandsgrößen* $\bar{Z}_1 = (\partial Z/\partial n_1)_{p,T,n_i \neq n_1}$, $\bar{Z}_2 = (\partial Z/\partial n_2)_{p,T,n_i \neq n_2} \ldots$ ein. $\bar{Z}_i$ entspricht der Änderung der Zustandsgröße Z der Mischphase bei Zugabe von 1 Mol des Stoffs X_i zu einer so großen Menge der Phase, daß sich dabei die Konzentrationen aller Stoffe praktisch nicht ändern.

Die partiellen molaren Größen hängen nur von dem Mengenverhältnis der Mischungskomponenten ab, nicht aber von der gesamten Menge der Mischphase. Sie stellen experimentell zugängige Rechengrößen dar, die jedoch nicht den wahren Zustandseigenschaften der einzelnen Komponenten innerhalb der Mischphase entsprechen, wie z. B. schon daraus hervorgeht, daß in der oben erwähnten Lösung von $MgSO_4$ in Wasser negative partielle molare Volumina $(\partial v/\partial n_{MgSO_4})_{n H_2O}$ auftreten. Diese ergeben sich daraus, daß das Wasser eine sehr sperrige Struktur hat (↑ 6.3.2.1.2.) und die Ionen Mg^{2+} und SO_4^{2-} infolge Solvatation eine Ordnung höherer Dichte der Teilchen hervorruft. Für jede beliebige extensive Eigenschaft gilt bei konstantem Druck und konstanter Temperatur die *Gleichung* von GIBBS und DUHEM $n_1\,d\bar{Z}_1 + n_2\,d\bar{Z}_2 + \ldots + n_k\,d\bar{Z}_k = \sum_{i=1}^{k} n_i\,d\bar{Z}_i = 0$, die sich unter Einführung der Molenbrüche $x_1, x_2 \ldots$ auch in der Form $x_1(\partial\bar{Z}_1/\partial x_j)_{T,p} + x_2(\partial\bar{Z}_2/\partial x_j)_{T,p} + \ldots + x_k(\partial\bar{Z}_k/\partial x_j)_{T,p} = 0$ schreiben läßt. Hierbei ist noch anzugeben, auf Kosten welchen Bestandteils der Mischung die Änderung des Molenbruchs x_j der Komponente X_j gehen soll. Für Zweistoffsysteme ist dies immer die andere Komponente; man erhält dann $x_1(\partial\bar{Z}_1/\partial x_1) = -x_2(\partial\bar{Z}_2/\partial x_1)$; $\bar{Z}_1$ und $\bar{Z}_2$ ändern sich bei Variation der Zusammensetzung in entgegengesetzter Richtung.

Mit Hilfe der partiellen molaren Größen und der Gibbs-Duhemschen Gleichungen lassen sich alle durch Änderungen der mengenmäßigen Zusammensetzung einer Mischphase auftretenden Zustandsänderungen in einfacher und eindeutiger Weise beschreiben.

6.1.2.2. Erster Hauptsatz, Energieerhaltungssatz. Jedes stoffliche System hat in jedem Zustand eine ganz bestimmte *innere Energie u*, die sich aus den inneratomaren und -molekularen Energien, aus den potentiellen Wechselwirkungsenergien und den kinetischen Energien aller Teilchen zusammensetzt. Um ein System vom Zustand 1 in einen Zustand 2 zu überführen, ist erfahrungsgemäß die Zufuhr oder Abgabe einer Wärme q oder einer mechanischen, elektrischen bzw. magnetischen Arbeit w oder von beiden notwendig.

Der *erste Hauptsatz* sagt aus:

Die innere Energie ist eine Zustandsfunktion $u = f(v, T, n_1, n_2, \ldots, n_k)$, man spricht von der *kalorischen Zustandsgleichung.*

Die Summe $q + w$ der zwischen einem System und seiner Umgebung ausgetauschten Wärme und Arbeit ist gleich der Änderung $u_2 - u_1$ der inneren Energie des Systems.

Dem System zugeführte Energiebeträge erhalten dabei ein positives, vom System abgegebene Energiebeträge ein negatives Vorzeichen. Das Verhältnis q/w hängt bei einer gegebenen Änderung der inneren Energie vom Wege der Zustandsänderung ab, z. B. wird $u_2 - u_1 = w$, wenn $q = 0$. Solche Vorgänge, bei denen kein Wärmeaustausch zwischen einem System und seiner Umgebung stattfindet, nennt man *adiabatische Vorgänge.*

6.1.2.2.1. *Enthalpie.* Neben der inneren Energie u verwendet man in der Thermodynamik häufig noch die *Enthalpie h* als Energiegröße. Sie ist durch die Gleichung $h = u + pv$ definiert und stellt daher ebenfalls eine Zustandsfunktion $h = f(p, T, n_1, n_2, \ldots, n_k)$ dar, für die man aber zweckmäßig nicht das Volumen v, sondern den Druck p als unabhängige Zustandsvariable wählt. Die Enthalpie eignet sich besonders zur Beschreibung isobarer Vorgänge ($dp = 0$), während bei isochoren Vorgängen ($dv = 0$) die innere Energie u den Vorzug verdient.

6.1.2.2.2. *Wärmekapazität, Molwärme.* Wird einem stofflichen System Wärme zugeführt, so erhöht sich seine Temperatur. Man nennt den Quotienten aus einer infinitesimal kleinen zugeführten Wärme dq und der dadurch verursachten infinitesimal kleinen Temperaturänderung dT die Wärmekapazität $c = dq/dT$ des Systems. Dividiert man die Wärmekapazität durch die Teilchenmenge n bzw. durch die Masse m des Systems, so gelangt man zur *molaren Wärmekapazität* oder *Molwärme* $C = c/n$ bzw. zur *spezifischen Wärmekapazität* oder *spezifischen Wärme* $C_{sp} = c/m$. Die bei konstantem Volumen ($dw = -p\,dv = 0$) zugeführte Wärme dq führt ausschließlich zu einer Erhöhung der inneren Energie du des Systems, so daß $c_v = (dq/dT)_v = (\partial u/\partial T)_v$. Die Wärmekapazität c_v bei konstantem Volumen entspricht somit dem Temperaturkoeffizienten der inneren Energie. Beim Erwärmen bei konstantem Druck wird wegen der Vergrößerung des Volumens zusätzliche Wärme zur Leistung von Volumenarbeit aufgewendet, so daß c_p stets größer ist als c_v. Die Wärmekapazität c_p bei konstantem Druck entspricht dem Temperaturkoeffizienten der Enthalpie: $c_p = (dq/dT)_p = (\partial h/\partial T)_p$. Allgemein gilt für einen reinen homogenen Stoff: $C_p - C_v = T(\partial p/\partial T)_V \cdot (\partial V/\partial T)_p = TV\gamma^2/\varkappa$ und speziell für ideale Gase (↑ 6.2.): $C_p - C_v = R$.

Die Molwärmen spielen in der Thermodynamik eine große Rolle. Sie sind experimentell zugängig und ermöglichen die Berechnung der Temperaturabhängigkeit der energetischen Zustandsgrößen der Stoffe; z. B. ergibt sich die Änderung der molaren inneren Energie bzw. Enthalpie eines Stoffs zwischen zwei Temperaturen T_1 und T_2 durch Integration zu $(U_2 - U_1)_V = \int_{T_1}^{T_2} C_v \, \mathrm{d}T$ bzw. $(H_2 - H_1)_p = \int_{T_1}^{T_2} C_p \, \mathrm{d}T$. Die Molwärmen sind stoffeigene Größen und hängen im allgemeinen stark von der Temperatur und dem Druck ab.

6.1.2.2.3. *Energieumsätze bei Stoffumwandlungen.* Die mit einem isotherm-isochoren bzw. isotherm-isobaren Stoffumwandlungsprozeß verbundene Änderung der inneren Energie $(\Delta u)_{v,T}$ bzw. Enthalpie $(\Delta h)_{p,T}$ eines Systems wird als *latente Wärme* bezeichnet. Sie entspricht der Wärme, die bei vollkommen irreversiblem Ablauf des Vorgangs zwischen dem System und seiner Umgebung ausgetauscht wird, ohne daß dabei eine Temperaturänderung auftritt. Je nach der Art des Vorgangs nennt man diese Wärme Schmelzwärme, Verdampfungswärme, Mischungswärme, Reaktionswärme chemischer Reaktionen u. a. Sind Δu bzw. Δh positiv, so muß man Wärme aus der Umgebung zuführen, damit die Temperatur während des Vorgangs konstant bleibt, der Vorgang heißt *endotherm*; bei negativem Δu bzw. Δh wird hingegen Wärme frei, der Vorgang heißt *exotherm.* Latente Wärmen können mit Hilfe von Kalorimetern exprimentell bestimmt werden.

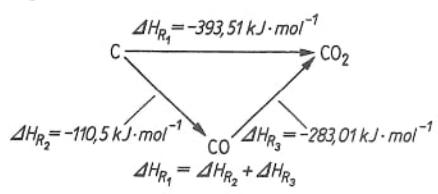


Abb. 6.1.2-3 Hesssches Gesetz am Beispiel der Reaktion $C + O_2 \rightarrow CO_2$ bei 25 °C, Standardbedingungen

Gesetz von HESS (1840): Da Δu und Δh Änderungen von Zustandsfunktionen sind, hängt ihr Wert nicht vom Wege der Zustandsänderung ab. Verläuft also z. B. eine chemische Reaktion einmal direkt und zum anderen über verschiedene Zwischenstufen von den Ausgangs- zu den Endstoffen, so sind die *Reaktionsenergien* $\Delta_R U$ bzw. *Reaktionsenthalpien* $\Delta_R H$ auf beiden Wegen gleich groß (Abb. 6.1.2-3).

Gesetz von KIRCHHOFF: Da der Temperaturkoeffizient der inneren Energie gleich der Wärmekapazität c_v und der der Enthalpie gleich c_p ist, ergibt sich die Temperaturabhängigkeit der latenten Wärme aus der Differenz der entsprechenden Wärmekapazitäten des End- und Ausgangszustands: $(\partial \Delta u/\partial T)_v = \Delta c_v$ bzw. $(\partial \Delta h/\partial T)_p = \Delta c_p$.

Zum Beispiel hat die chemische Reaktion $C + 2\,H_2 \rightarrow CH_4$ bei 298,15 K (25 °C) und 10^5 Pa die Reaktionsenthalpie $\Delta_R H = -74{,}90\ \mathrm{kJmol^{-1}}$. Die Molwärmen der Reaktanten betragen:

C (Graphit): $C_p(T) = (16{,}87 + 4{,}77 \cdot 10^{-3}\,T - 8{,}54 \cdot 10^5\,T^{-2})\ \mathrm{JK^{-1}mol^{-1}}$

H_2 (Gas): $C_p(T) = (27{,}30 + 3{,}27 \cdot 10^{-3}\,T + 0{,}50 \cdot 10^5\,T^{-2})\ \mathrm{JK^{-1}mol^{-1}}$

CH_4 (Gas): $C_p(T) = (23{,}66 + 47{,}90 \cdot 10^{-3}\,T - 1{,}93 \cdot 10^5\,T^{-2})\ \mathrm{JK^{-1}mol^{-1}}$

und damit $(\partial \Delta_R H/\partial T)_p = \Delta_R C_p = C_p(CH_4) - C_p(C) - 2\,C_p(H_2) = (-47{,}81 + 36{,}59 \cdot 10^{-3}\,T + 5{,}61 \cdot 10^5\,T^{-2})\ \mathrm{JK^{-1}mol^{-1}}$.

Die Integration dieser Gleichung liefert $\Delta_R H(T) = (-47{,}81\,T + 18{,}30 \cdot 10^{-3}\,T^2 - 5{,}61 \cdot 10^5\,T^{-1} - 60390)\ \mathrm{Jmol^{-1}}$.

6.1.2.2.4. *Bildungswärmen.* Aus den bei der Verbrennung von Elementen und chemischen Verbindungen in reinem Sauerstoff auftretenden *Verbrennungswärmen* kann man mit Hilfe des Hessschen Gesetzes die meist nicht direkt meßbaren Bildungswärmen berechnen, die bei der Bildung einer chemischen Verbindung aus ihren Elementen bei konstantem Druck auftreten würden. Diese werden als *Bildungsenthalpie* $\Delta_B H$ bezeichnet und für 1 Mol der Verbindung angegeben.

Da Reaktionswärmen von der Zusammensetzung des Reaktionssystems abhängen, setzt man für den Vergleich z. B. von Bildungsenthalpien fest, daß alle Reaktionsteilnehmer in bestimmten *Standardzuständen* vorliegen sollen. Bei Gasen ist dies der Zustand des idealen Gases bei $1{,}01325 \cdot 10^5$ Pa, bei flüssigen und festen Stoffen der Zustand der reinen Stoffe und bei gelösten Stoffen ihr Zustand bei unendlicher Verdünnung. Man setzt weiterhin fest, daß die Bildungsenthalpie der Elemente in ihrer stabilen Form und in ihren Standardzuständen gleich Null sein soll. Die *Standardbildungsenthalpien* $\Delta_B H^\ominus$ zahlreicher Stoffe sind gewöhnlich für 25 °C tabelliert. Aus ihnen lassen sich mit Hilfe der Gesetze von HESS und KIRCHHOFF die Reaktionsenthalpien beliebiger Reaktionen unter Standardbedingungen errechnen.

6.1.2.3. Zweiter Hauptsatz, Entropiesatz. Die Erfahrung lehrt, daß alle Vorgänge in der Natur nur in ganz bestimmten Richtungen von selbst ablaufen. Jedes stoffliche System nimmt schließlich den unter den jeweiligen Bedingungen stabilsten Zustand, den Gleichgewichtszustand, ein. Ist dieser erreicht, so kehrt es niemals wieder von selbst in einen früheren Zustand zurück, z. B. tritt in einem System mit von Ort zu Ort verschiedener Temperatur von selbst ein Temperaturausgleich ein oder zwei verschiedene Gase vermischen sich ohne äußeres

Zutun so lange, bis alle Konzentrationsunterschiede ausgeglichen sind.

6.1.2.3.1. *Entropie*. Als Maß für die Richtung derartiger spontan verlaufender irreversibler Zustandsänderungen dient nach CLAUSIUS (1854) die Entropie s. Sie ist ebenso wie die innere Energie u oder Enthalpie h eine thermodynamische Zustandsfunktion $s = f(v$ bzw. $p, T, n_1, n_2, \ldots, n_k)$. Sie steht nach der Gleichung $s = s_0 + \int_0^T dq_{rev}/T$ in unmittelbarem Zusammenhang mit der zwischen dem betrachteten System und seiner Umgebung ausgetauschten Wärme q, wenn man das System auf reversiblem Wege vom absoluten Nullpunkt in den jeweils betrachteten Zustand der Temperatur T bringt. Bei der Berechnung dieses Integrals hat man jede Wärme dq_{rev} durch die Temperatur T, bei der sie ausgetauscht wurde, zu dividieren. s_0 ist die Entropie am absoluten Nullpunkt. Die *Nullpunktentropie* s_0 von chemisch homogenen Stoffen, die sich im inneren Gleichgewicht befinden, ist nach PLANCK Null. Bei Gläsern oder Mischkristallen, die bei $T = 0$ nicht im inneren Gleichgewicht sind, ist jedoch $s_0 \neq 0$. Diese Aussagen stellen eine Erweiterung des Wärmetheorems von NERNST dar, nach dem für jeden Prozeß in einem geschlossenen System, an dem nur Phasen im inneren Gleichgewicht beteiligt sind, $\lim_{T \to 0} \Delta s = 0$ gilt. Die Entropie ist ein Maß für die Stabilität eines Systems. Sie wird anschaulich, wenn man sie mit der *Zustandswahrscheinlichkeit* (↑ 2.4.2.3.) eines Systems verbindet. Der Zustand maximaler Stabilität ist auch der Zustand größter Wahrscheinlichkeit, den jedes System einzunehmen sucht. Die Wahrscheinlichkeit ist um so größer, je größer die Zahl der Anordnungsmöglichkeiten der Bausteine (Mikrozustände) des Systems bzw. je größer der Unordnungszustand des Systems ist. Ein Maß für die Zahl der Anordnungsmöglichkeiten ist die „*thermodynamische Wahrscheinlichkeit*" W (↑ 6.1.3.1.). Nach BOLTZMANN gilt $s = k_B \ln W$ ($k_B = 1{,}3803 \cdot 10^{-23}$JK^{-1} Boltzmannkonstante). Je ungeordneter ein System ist, um so größer ist seine Entropie und um so stabiler ist sein Zustand.

Der 2. Hauptsatz der Thermodynamik besagt, für ein abgeschlossenes System:

Bei allen von selbst verlaufenden, d. h. irreversiblen Vorgängen, tritt eine Entropiezunahme auf, $\Delta s_{irr} > 0$.

Zustandsänderungen können nur so lange spontan vor sich gehen, wie die Entropie noch zunehmen kann. Die Entropie strebt einem Maximum zu, und das System befindet sich dann im Gleichgewicht, wenn dieses Maximum erreicht ist. Bei allen Vorgängen, die reversibel, d. h. im Gleichgewichtszustand, verlaufen, bleibt die Entropie konstant, $\Delta s_{rev} = 0$.

Für eine differentielle Zustandsänderung gilt somit $ds \geqq dq/T$; dabei gilt die Ungleichheit bei irreversiblen und die Gleichheit bei reversiblen Vorgängen.

6.1.2.3.2. *Freie Energie und freie Enthalpie*. Die vorstehenden Aussagen des 2. Hauptsatzes (s. a. 6.1.2.3.1.) gelten ausschließlich für abgeschlossene Systeme, deren innere Energie während des Ablaufs irgendwelcher Vorgänge wegen des Energieerhaltungssatzes konstant bleiben muß. Bei nicht abgeschlossenen Systemen, die also im Energieaustausch mit der Umgebung stehen, kann die Tendenz nach einem maximalen Wert der Entropie durch die nach einem minimalen Wert der inneren Energie überlagert werden. Durch die Kombination der Gleichung $du = dq + dw$ für den ersten Hauptsatz mit der Gleichung $dq \leqq T\,ds$ für den zweiten erhält man $du - T\,ds \leqq dw$ bzw. für einen konstanten Wert der Temperatur $d(u - Ts)_T \leqq dw$. In dieser Beziehung entspricht das Gleichheitszeichen reversiblen, die Ungleichheit dagegen irreversiblen, d. h. von selbst verlaufenden, Zustandsänderungen.

Die Zustandsfunktion $f = u - Ts$ wird als *freie Energie* bezeichnet. Sie entspricht demjenigen Anteil an der inneren Energie eines Systems, der bei konstanter Temperatur maximal in Arbeit umgewandelt werden kann, der für eine Arbeitsleistung „frei" verfügbar ist. Für Vorgänge mit $dw = 0$, die bei konstanter Temperatur und konstantem Volumen vor sich gehen, wird $df = d(u - Ts)_{T,v} \leqq 0$. Diese Beziehung sagt aus, daß Vorgänge in nicht abgeschlossenen Systemen bei konstanter Temperatur und konstantem Volumen nicht in Richtung auf ein Maximum der Entropie, sondern nach einem Minimum der freien Energie ablaufen. Ist dieses Minimum der freien Energie erreicht, ist das System im Gleichgewicht.

Für Zustandsänderungen bei konstanter Temperatur und konstantem Druck, die in der Natur und in der Technik weit häufiger auftreten, gilt analog $dg = d(h - Ts)_{T,p} \leqq 0$; dabei wird die Zustandsfunktion $g = h - Ts$ als *freie Enthalpie* bezeichnet. Die Vorgänge laufen in diesem Fall in Richtung nach einem Minimum der freien Enthalpie ab, das im Gleichgewichtszustand erreicht ist.

Bei konstanter Temperatur und konstantem Volumen bzw. konstantem Druck können Transportvorgänge, Phasenumwandlungen, chemische Reaktionen u. a. nur dann ablaufen, wenn die Änderung der freien Energie $\Delta f = \Delta u - T\Delta s$ für konstantes T und v bzw. die der freien Enthalpie $\Delta g = \Delta h - T\Delta s$ für konstantes T und p einen negativen Wert hat.

Diese beiden Gleichungen werden als *Gleichungen* von GIBBS und HELMHOLTZ bezeichnet. In ihnen kommen deutlich die beiden einander entgegenwirkenden Tendenzen zum Ausdruck,

die der inneren Energie bzw. der Enthalpie nach einem minimalen und die der Entropie nach einem maximalen Wert. Da ein Vorgang um so leichter eintreten kann, je negativer Δf bzw. Δg ist, nennt man $-\Delta f$ bzw. $-\Delta g$ häufig auch *Affinität* oder Triebkraft eines Vorgangs. Sie ist um so größer, je negativer die Größe Δu bzw. Δh und je positiver Δs ist, mit anderen Worten, je exothermer der Vorgang und je größer die Entropiezunahme dabei ist. Nach den Gibbs-Helmholtzschen Gleichungen kann somit die Richtung eines Vorgangs bestimmt werden; z. B. kann die Oxydation des Kohlenstoffs entweder nach der Gleichung $C + O_2 \rightarrow CO_2$ zum Kohlendioxid oder nach der Gleichung $2\,C + O_2 \rightarrow 2\,CO$ zum Kohlenmonoxid führen. Bei dem ersten Vorgang wird viel mehr Wärme frei, nimmt die Enthalpie h viel stärker ab, als beim zweiten. Demgegenüber ändert sich die Entropie beim ersten praktisch überhaupt nicht, während sie beim zweiten stark zunimmt. Bestünde nur die Tendenz, eine minimale Enthalpie zu erreichen, würde aller Kohlenstoff zu CO_2 verbrennen, wäre andererseits nur die Tendenz vorhanden, eine maximale Entropie zu erreichen, würde bei der Verbrennung ausschließlich CO entstehen. In Wirklichkeit entsteht eine Mischung beider Gase. Bei niedrigen Temperaturen überwiegt jedoch der Energie- oder Enthalpieeffekt, und Kohlenstoff verbrennt dann hauptsächlich zu CO_2; bei hohen Temperaturen bekommt wegen der Multiplikation mit T der Entropieeffekt das Übergewicht, und es entsteht hauptsächlich CO.

6.1.2.4. Thermodynamische Gleichgewichtsbedingung. Da isotherm-isochore bzw. isotherm-isobare Vorgänge nur dann von selbst verlaufen können, wenn $(\Delta f)_{T,v}$ bzw. $(\Delta g)_{T,p}$ negativ ist, laufen Vorgänge, für die diese Größen positive Werte haben, freiwillig in der umgekehrten Richtung ab. Gilt für einen Vorgang $(\Delta f)_{T,v} = 0$ bzw. $(\Delta g)_{T,p} = 0$, so kann er weder in der einen noch in der anderen Richtung von selbst verlaufen, das betrachtete System befindet sich im Gleichgewicht.

Die *allgemeinen thermodynamischen Gleichgewichtsbedingungen* lauten $\Delta f = 0$ für konstantes T und v und $\Delta g = 0$ für konstantes T und p.

Für die Ermittlung eines Gleichgewichtszustands muß die Abhängigkeit der Größen Δf bzw. Δg von allen Zustandsvariablen $v, p, T, n_1, n_2, \ldots, n_k$ berücksichtigt werden. Für differentielle Zustandsänderungen in einer Phase mit k verschiedenen Stoffen X_i liefert der 1. Hauptsatz bei alleiniger Volumenarbeit $du = dw + dq = -p\,dv + dq$, während nach dem 2. Hauptsatz $dq = T\,ds$ gilt; unter zusätzlicher Beachtung der Abhängigkeit von den Molzahlen n_i der Komponenten X_i erhält man daraus $df = -p\,dv - s\,dT + \sum_{i=1}^{k} (\partial f/\partial n_i)\,dn_i$ bzw. $dg = v\,dp - s\,dT + \sum_{i=1}^{k} (\partial g/\partial n_i)\,dn_i$. Die Größen $(\partial f/\partial n_j)_{T,v,n_i \neq n_j} = (\partial g/\partial n_j)_{T,p,n_i \neq n_j} = \mu_j$ sind die *partielle molare freie Energie* bzw. *Enthalpie* des Stoffs X_j innerhalb der betrachteten Phase und werden als *chemische Potentiale* bezeichnet. Sie entsprechen der Änderung der freien Energie bzw. der freien Enthalpie der Phase bei Zugabe von 1 Mol des Stoffs X_j zu einer so großen Menge der Phase, daß sich dabei die Konzentration der Stoffe praktisch nicht ändert. Da im Gleichgewicht $(df)_{T,v} = (dg)_{T,p} = 0$ ist, erhält man somit als eine andere Form der allgemeinen Gleichgewichtsbedingung für jede beliebige Phase $\sum_{i=1}^{k} \mu_i\,dn_i = 0$. Die thermodynamischen Gleichgewichtsbedingungen gelten grundsätzlich für alle Arten stofflicher Gleichgewichte. Von besonderer praktischer Bedeutung sind die Phasengleichgewichte (↑ 6.1.2.4.2.), die chemischen Gleichgewichte (↑ 6.1.2.4.3.) und die Grenzflächengleichgewichte (↑ 6.6.).

6.1.2.4.1. *Chemisches Potential.* Das chemische Potential μ_i eines Stoffs X_i in einer Mischphase hängt von der Temperatur, dem Druck und der Zusammensetzung der Mischphase ab. Für die Druck- und Temperaturabhängigkeit folgt aus dem 1. und 2. Hauptsatz $(\partial\mu_i/\partial p)_{T,n_i} = \bar{V}_i$ und $(\partial\mu_i/\partial T)_{p,n_i} = -\bar{S}_i$, wobei $\bar{V}_i$ das partielle molare Volumen und $\bar{S}_i$ die partielle molare Entropie des Stoffs X_i sind. Daraus und aus dem Gasgesetz $pV = RT$ ergibt sich $d\mu_i = V\,dp = RT\,dp/p$, d. h. für die Druckabhängigkeit gilt im Falle von *idealen Gasmischungen* $\mu_i = \mu_i^{\ominus} + RT \ln(p_i/p_t)$, wenn p_i der Partialdruck des Gases X_i und p_t der Normdruck von $1{,}0132 \cdot 10^5$ Pa ist. Das *chemische Standardpotential* $\mu_i^{\ominus}$ entspricht dabei dem chemischen Potential des idealen Gases im Standardzustand bei $\approx 10^5$ Pa Druck und ist nur noch von der Temperatur abhängig. Um auch für *reale Gasgemische* eine analoge einfache Beziehung $\mu_i = \mu_i^{\ominus} + RT \ln(p_i^*/p_t)$ zu erhalten, führt man anstelle des wirklichen Partialdrucks p_i eine neue Druckgröße p_i^* ein, die *Fugazität*, die mit dem wahren Druck identisch wird, wenn dieser so gering wird, daß das Gas in den idealen Zustand übergeht: $\lim p^*/p = 1$ für $p \rightarrow 0$. Auch bei *kondensierten* flüssigen oder festen *Mischphasen* kann man zwischen idealen und realen Mischungen unterscheiden (↑ 6.3.3.3.). Bei idealen Mischungen ist die Wechselwirkungsenergie zwischen zwei verschiedenen Teilchen gleich dem arithmetischen Mittel der Wechselwirkungsenergien zwischen den gleichartigen Teilchen. In diesem Falle beträgt die Zusammensetzungsabhängigkeit des chemischen Potentials $\mu_i = \mu_i^{\ominus} + RT \ln x_i$, wenn x_i der Molenbruch des Stoffs X_i in der Mischung ist. Bei *realen kondensierten Mischphasen* führt man analog wie bei realen

Gasmischungen einen korrigierten Molenbruch $a_i = f_i x_i$ ein, den man als *Aktivität* des Stoffs X_i bezeichnet. Der Korrekturfaktor f_i, der Aktivitätskoeffizient, berücksichtigt sämtliche durch das nichtideale Verhalten einer Mischung bedingten Abweichungen von den idealen Gesetzen und muß in der Regel experimentell bestimmt werden. Für das chemische Potential des Stoffs X_i erhält man somit $\mu_i = \mu_i^{\ominus} + RT \ln a_i$. Für den reinen Stoff X_i wird $x_i = f_i = a_i = 1$ gesetzt. Daher entspricht $\mu_i^{\ominus}$ dem chemischen Potential des reinen Stoffs X_i; es wird als *chemisches Standardpotential* bezeichnet und ist nur noch vom Druck und von der Temperatur abhängig.

Bei flüssigen oder festen Lösungen, in denen der gelöste Stoff in relativ kleinen Mengen gegenüber dem Lösungsmittel vorhanden ist, wählt man nun als fiktiven Standardzustand für den gelösten Stoff häufig seinen Zustand in der ideal verdünnten Lösung. Der Aktivitätskoeffizient nähert sich dann mit abnehmender Konzentration dem Wert 1; $\lim f_i = 1$ für $x_i \to 0$.

6.1.2.4.2. *Phasengleichgewichte* (↑ 6.3., 6.4.). Besteht ein System aus mehreren Phasen I, II, ... usw., so ist das Phasengleichgewicht dadurch gekennzeichnet, daß jeder Stoff in jeder Phase in einer bestimmten, zeitlich unveränderlichen Konzentration vorhanden ist. Für jede Phase muß Summe $\mu_i \, dn_i = 0$ gelten, und die Gleichgewichtsbedingung für das ganze System lautet:

$$\sum_{i=1}^{k} \mu_i^{\mathrm{I}} \, dn_i^{\mathrm{I}} + \sum_{i=1}^{k} \mu_i^{\mathrm{II}} \, dn_i^{\mathrm{II}} + \ldots = 0.$$

Bei einem Übergang von dn_i Molen eines Stoffs X_i aus der Phase I in die Phase II ist $dn_i^{\mathrm{I}} = -dn_i^{\mathrm{II}}$, und es folgt $\mu_i^{\mathrm{I}} = \mu_i^{\mathrm{II}} = \ldots$

Im Phasengleichgewicht muß das chemische Potential eines jeden Stoffes in jeder Phase denselben Wert haben. Man nennt diese Aussage die *Koexistenzbedingung*.

Ebenso wie elektrische Potentialdifferenzen zu einem elektrischen Strom führen, bewirken Unterschiede im chemischen Potential den freiwilligen Übergang des betreffenden Stoffs von einer Phase in die andere.

6.1.2.4.3. *Chemische Gleichgewichte.* Der stoffliche Umsatz beim Ablauf einer chemischen Reaktion kann durch eine *Bruttoreaktionsgleichung* beschrieben werden. Die Reaktionsgleichung $2\,H_2 + O_2 \to 2\,H_2O$ besagt qualitativ, daß sich aus Wasserstoff und Sauerstoff Wasser bildet, und quantitativ, daß jeweils 2 Moleküle bzw. 2 Mole H_2 mit 1 Molekül bzw. 1 Mol O_2 gerade 2 Moleküle bzw. 2 Mole H_2O ergeben. Bei einer beliebigen Reaktion $\nu_A A + \nu_B B \to \nu_C C + \nu_D D$ sind die Stoffe A und B die *Ausgangsstoffe* oder *Reaktanten* und die Stoffe C und D die *Endstoffe* oder *Reaktionsprodukte*. Die Zahlen ν nennt man die Stöchiometriezahlen. Als quantitatives Maß für den jeweiligen Umsatz dient die *Reaktionslaufzahl* λ. Sie gibt an, welcher Bruchteil bzw. welches Vielfache eines Formelumsatzes stattgefunden hat. Ein infinitesimaler Umsatz $d\lambda$ entspricht Änderungen der Molzahlen der Reaktionsteilnehmer um $dn_A = -\nu_A \, d\lambda$, $dn_B = -\nu_B \, d\lambda$, $dn_C = +\nu_C \, d\lambda$ und $dn_D = +\nu_D \, d\lambda$. Damit folgt aus der allgemeinen Gleichgewichtsbedingung $(df)_{T,v} = (dg)_{T,p} = \nu_C \mu_C \, d\lambda + \nu_D \mu_D \, d\lambda - \nu_A \mu_A \, d\lambda - \nu_B \mu_B \, d\lambda = 0$ oder $\nu_A \mu_A + \nu_B \mu_B = \nu_C \mu_C + \nu_D \mu_D$. Unter Berücksichtigung der Konzentrationsabhängigkeit des chemischen Potentials $\mu_i = \mu_i^{\ominus} + RT \ln a_i$ erhält man daraus für konstante Temperatur und konstanten Druck, d. h. konstante Werte von $\mu_i^{\ominus}$, die Gleichung $RT(\nu_A \ln a_A + \nu_B \ln a_B - \nu_C \ln a_C - \nu_D \ln a_D) = \nu_C \mu_C^{\ominus} + \nu_D \mu_D^{\ominus} - \nu_A \mu_A^{\ominus} - \nu_B \mu_B^{\ominus} = -RT \ln K$ oder $a_C^{\nu_C} \cdot a_D^{\nu_D} / (a_A^{\nu_A} \cdot a_B^{\nu_B}) = K$. Diese als *Massenwirkungsgesetz* bezeichnete Beziehung besagt, daß bei konstantem p und T im chemischen Gleichgewicht stets ein genau bestimmtes Verhältnis der Aktivitäten der Reaktionsteilnehmer zueinander vorliegen muß. Die Konstante K wird *thermodynamische Gleichgewichtskonstante* genannt. Sie ist nur von p und T abhängig.

Je größer K ist, um so größer ist der mögliche Umsetzungsgrad, d. h. die erzielbare Ausbeute an Reaktionsprodukten. Bei Reaktionen in ideal verdünnten Lösungen oder idealen Gasgemischen sind die Aktivitäten a_i gleich den Konzentrationen c_i bzw. den Partialdrücken p_i, so daß das Massenwirkungsgesetz die einfachere Form $c_C^{\nu_C} \cdot c_D^{\nu_D} / (c_A^{\nu_A} \cdot c_B^{\nu_B}) = K_c$ bzw. $p_C^{\nu_C} \cdot p_D^{\nu_D} / (p_A^{\nu_A} \cdot p_B^{\nu_B}) = K_p$ annimmt. In den praktisch meist vorliegenden realen Reaktionsgemischen sind K_c und K_p keine konstanten Größen.

Liegt bei heterogenen Reaktionen ein Reaktionsteilnehmer als reiner Stoff vor, so tritt dessen Aktivität im Massenwirkungsgesetz überhaupt nicht auf, da die Aktivität reiner Stoffe definitionsgemäß stets gleich 1 ist.

Für die Temperaturabhängigkeit der Gleichgewichtskonstanten bei konstantem Druck gilt die *van't-Hoffsche Reaktionsisobare* $(\partial \ln K / \partial T)_p = \Delta_R H^{\ominus} / RT^2$, wobei $\Delta_R H^{\ominus}$ die Enthalpieänderung pro Formelumsatz im Standardzustand, die Standard-Reaktionsenthalpie, bedeutet. Man gewinnt diese Beziehung aus $\nu_C \mu_C^{\ominus} + \nu_D \mu_D^{\ominus} - \nu_A \mu_A^{\ominus} - \nu_B \mu_B^{\ominus} = \Delta_R G^{\ominus} = -RT \ln K$ durch Differentiation nach T, wenn man beachtet, daß der Temperaturkoeffizient der *freien Standardreaktionsenthalpie* $(\partial \Delta_R G^{\ominus} / \partial T)_p = -\Delta_R S^{\ominus}$ der *negativen Standardreaktionsentropie* gleich ist (↑ 6.1.2.4.1.) und nach GIBBS-HELMHOLTZ $\Delta_R G^{\ominus} + T \Delta_R S^{\ominus} = \Delta_R H^{\ominus}$ gilt. Die Druckabhängigkeit von K bei konstanter Temperatur ergibt sich zu $(\partial \ln K / \partial p)_T = -\Delta_R V^{\ominus} / RT$, wobei $\Delta_R V^{\ominus}$ die Volumenänderung pro Formelumsatz im Standardzustand, das *Standardreaktionsvolumen*, angibt.

6.1.3. Statistische Thermodynamik

Die moderne Physik und Chemie versuchen, die atomare und molekulare Natur der stofflichen Materie zu verstehen. Sie benutzen dazu in erster Linie die Sprache der klassischen Mechanik, der Quantenmechanik und der Statistik. Zwischen den mikroskopischen und den in der phänomenologischen Thermodynamik betrachteten makroskopischen Eigenschaften muß ein Zusammenhang bestehen. Wenn man ihn kennt, wird es möglich, makroskopische Eigenschaften aus den Eigenschaften der korpuskularen Bausteine der Materie und ihren Wechselwirkungen zu erklären oder umgekehrt aus makroskopischen Eigenschaften gewisse Rückschlüsse auf die Natur der Teilchen und ihrer Anhäufungen zu ziehen. Diese Zielstellung ist Aufgabe der statistischen Thermodynamik (↑ 2.4.).

6.1.3.1. Zusammenhang zwischen statistischen und thermodynamischen Größen

6.1.3.1.1. *Innere Energie.* Bei einem *abgeschlossenen System*, das mit seiner Umgebung weder Energie noch Materie austauschen kann, muß die innere Energie u dauernd gleich dem gesamten inneren Energievorrat E sein, der in dem System einmal gespeichert worden ist (↑ 6.1.2.2.). In einem *geschlossenen Teilsystem*, das mit seiner Umgebung Energie, aber keine Materie austauscht, kann hingegen die Energie verschiedene Werte E_n annehmen. Der für einen thermodynamischen Zustand scharf definierte Wert der inneren Energie u kann dann als zeitlicher Mittelwert über die verschiedenen vom System wirklich angenommenen Energiewerte verstanden werden. Im Gleichgewicht ergibt sich somit die innere Energie $u = \sum_n w_n E_n$, wobei w_n die Wahrscheinlichkeit der Quantenzustände n bedeutet. Bei einem makroskopischen System mit einer sehr großen Anzahl von Teilchen und Freiheitsgraden sind die zu erwartenden Abweichungen des Meßwerts vom Mittelwert vieler Messungen selbst bei sehr kurzen Beobachtungszeiten außerordentlich klein.
Führt man für $w_n(E_n)$ den Ausdruck 2.4.(7) ein, so erhält man

$$(1) \quad u = 1/Z \sum_n \exp[-E_n/(k_B T)]\, E_n = k_B T^2 \,\mathrm{d} \ln Z/\mathrm{d}T.$$

Z ist die Zustandssumme (↑ 2.4.4.2.), k_B die Boltzmannkonstante und T die absolute Temperatur.

6.1.3.1.2. *Entropie.* Alle aus der phänomenologischen Thermodynamik bekannten Eigenschaften der Entropie (↑ 6.1.2.3.1.) werden durch die Größe

$$(2) \quad s = -k \sum_n w_n \ln w_n = u/T + k \ln Z$$

ichtig wiedergegeben (↑ 2.4.1.6.).

Jede Veränderung einer makroskopischen Eigenschaft eines Systems muß unter gleichbleibenden äußeren Bedingungen auf einer Veränderung der Wahrscheinlichkeiten w_n der Quantenzustände n beruhen, weil die Eigenschaften der Quantenzustände selbst konstant bleiben. Eine Änderung von s läßt sich daher ausdrücken durch $\mathrm{d}s = -k \sum_n (\ln w_n + 1)\,\mathrm{d}w_n$.
In einem *abgeschlossenen System* haben alle möglichen Quantenzustände der Gesamtzahl W die gleiche Energie E und damit die gleiche Wahrscheinlichkeit $w_n = 1/W$ (↑ 2.4.4.2.). Da die Summe aller Wahrscheinlichkeiten stets Eins ist, folgt somit $\sum_n \mathrm{d}w_n = 0$ und $\mathrm{d}s = 0$. Dies bedeutet, daß mit den im Gleichgewicht geltenden Wahrscheinlichkeiten die Größe s einen Extremwert annimmt. Dieser entspricht einem Maximum, da die zweite Ableitung von s nach w_n negativ ist. Die Größe s verhält sich somit tatsächlich wie die Entropie der phänomenologischen Thermodynamik.
Im Gleichgewicht eines abgeschlossenen Systems ist $s = k \ln W$; die Entropie steigt logarithmisch mit der Zahl der möglichen Quantenzustände, d. h. mit der Zahl der Realisierungsmöglichkeiten W für den betrachteten Energiezustand E, an. Die Größe W wird daher oft auch als *thermodynamische Wahrscheinlichkeit* bezeichnet.
Hat ein *geschlossenes System* konstanten Volumens, das mit seiner Umgebung keine Arbeit austauscht, m zugängliche Quantenzustände, so wäre seine Entropie maximal, wenn wegen der thermodynamischen Beziehung $\mathrm{d}u = \sum_n E_n \,\mathrm{d}w_n = T\,\mathrm{d}s$ die Bedingung $\mathrm{d}s = \sum_1^m E_n/T \,\mathrm{d}w_n = 0$ erfüllt würde. Dies kann nur für unendlich hohe Temperatur allgemein gültig sein. Dann aber sind alle w_n gemäß 2.4.(7) gleich $1/m$. Dies bedeutet, daß das System alle Quantenzustände mit gleicher Wahrscheinlichkeit besetzt und nicht wie bei endlichen Temperaturen die energetisch tiefer liegenden bevorzugt. Je bevorzugter einzelne Quantenzustände werden, je geordneter das System wird, um so kleiner wird seine Entropie. Beim absoluten Nullpunkt sind nur die W_1 Quantenzustände der niedrigsten Energie E_1 besetzt. Die Entropie wird dann $s = k \ln W_1$. Für $W_1 = 1$ wird sie Null, womit auch das Nernstsche Wärmetheorem (↑ 6.1.2.3.1.) erfüllt ist.
Vergleicht man die Entropie zweier verschiedener Systeme bei der gleichen Temperatur, so ist im allgemeinen die Entropie desjenigen Systems größer, das die größere Zahl von Quantenzuständen in einem Energiebereich bis zu einigen „kT" hat. Deshalb wird z. B. ein Stoff, der aus sehr flexiblen Teilchen besteht, eine größere molare Entropie haben als ein anderer, dessen Teilchen aus den gleichen Atomen bestehen, bei

dem aber infolge größerer Starrheit der Konfiguration ein Teil der Quantenzustände zu höheren Energien verschoben ist. In diesem Sinne muß die landläufige Aussage, daß die Entropie ein Maß für die „Unordnung" sei, verstanden werden.

6.1.3.1.3. *Weitere thermodynamische Funktionen* (↑ 6.1.2.). Der statistische Ausdruck für die *freie Energie* $f = u - Ts$ wird mit 6.1.(1) und 6.1.(2) $f = -kT \ln Z$. Damit läßt sich auch die *Entropie* $s = -(\partial f/\partial T)_v$ allein aus der Zustandssumme ableiten:

(3) $s = k[\ln Z + T(\partial \ln Z/\partial T)_v]$.

Weiter folgt für den *Druck*

(4) $p = -(\partial f/\partial v)_T = kT(\partial \ln Z/\partial v)_T$,

für die *Enthalpie*

$$h = u + pv = kT[T(\partial \ln Z/\partial T)_v + v(\partial \ln Z/\partial v)_T],$$

für die *freie Enthalpie*

$$g = h - Ts = kT[v(\partial \ln Z/\partial v)_T - \ln Z].$$

6.1.3.2. Berechnung von Zustandssummen. Die Berechnung der Zustandssumme eines makroskopischen Systems ist im allgemeinen schwer, weil sie die Kenntnis der vielen möglichen Energiezustände und ihrer Entartungsgrade voraussetzt. Sie läßt sich aber häufig vereinfachen, wenn das geschlossene System in unabhängige Untersysteme aufteilbar ist. Unter „unabhängig" wird hier verstanden, daß sich der Zustand eines Untersystems unabhängig vom momentanen Zustand eines anderen beschreiben läßt. Die Energie des Gesamtsystems ist dann gleich der Summe der Energien der Untersysteme. Zum Beispiel können in einem verdünnten Gas die translatorischen Freiheitsgrade eines Teilchens zu einem Untersystem zusammengefaßt werden und die Schwingungsfreiheitsgrade zu einem anderen, davon unabhängigen. Mit etwas weniger guter Näherung lassen sich auch die Rotationsfreiheitsgrade zu einem von dem Untersystem der Schwingungen unabhängigen Untersystem zusammenfassen.

Damit alle Untersysteme miteinander ins thermische Gleichgewicht kommen können, ist aber eine minimale Kopplung notwendig. Zwischen Translation und Schwingung wird sie z. B. durch Stöße zwischen den Teilchen vermittelt.

6.1.3.2.1. *Unterscheidbare, unabhängige Untersysteme.* Das i-te von m Untersystemen habe die Quantenzustände $n(i)$ mit den Energien $E_{n(i)}$. Wenn man es als geschlossenes Teilsystem im Gleichgewicht bei der Temperatur T betrachtet, so ist $w_{n(i)} = (1/Z_i) \exp[-E_{n(i)}/(kT)]$, nach 2.4.(7). Wenn sich in einem größeren geschlossenen System zu gleicher Zeit das Untersystem 1 in einem Quantenzustand mit der Energie $E_{n(1)}$, das Untersystem 2 in einem Quantenzustand mit der Energie $E_{n(2)}$ usw. befinden, ist die Gesamtenergie $E = \sum_{i=1}^{m} E_{n(i)}$. Die Wahrscheinlichkeit für diesen Zustand ist gleich dem Produkt der Einzelwahrscheinlichkeiten $w_{n(i)}$, und es folgt $Z = \prod_{i=1}^{m} Z_i$.

Die Zustandssumme eines geschlossenen Systems ist gleich dem Produkt der Zustandssummen seiner unterscheidbaren unabhängigen Untersysteme.

6.1.3.2.2. *Nicht unterscheidbare, unabhängige Untersysteme.* Als Beispiel kann ein reines ideales Gas (↑ 6.2.1.) dienen, wobei seine Teilchen als Untersysteme betrachtet werden. Da diese nicht unterscheidbar sind, kann man nicht feststellen, ob ein erstes Untersystem in einem Quantenzustand n und ein zweites in einem Quantenzustand k ist oder umgekehrt. Bei einer Vertauschung dieser beiden gleichen Untersysteme bleibt das System insgesamt im gleichen Quantenzustand.

Unter der Voraussetzung, daß von m Untersystemen jedes in einem anderen Quantenzustand ist, gibt es $m!$ Möglichkeiten, die Quantenzustände auf die Untersysteme zu verteilen. Wegen ihrer Nichtunterscheidbarkeit zählen alle Möglichkeiten zusammen als *ein* Quantenzustand, und es wird $Z = (1/m!)Z_i^m$. Die Voraussetzung, daß alle Untersysteme verschiedene Quantenzustände haben, trifft insbesondere für tiefe Temperaturen und bei vielen identischen Untersystemen oft nicht mehr zu, weil dann nur die energetisch niedrigsten Quantenzustände eingenommen werden. In diesem Fall muß man berücksichtigen, daß nach den Postulaten der Quantenmechanik (↑ 2.5.) bestimmte Kombinationen von Quantenzuständen der Untersysteme nicht vorkommen können oder daß verschiedene solcher Kombinationen zusammen als ein Quantenzustand zu zählen sind.

6.1.3.3. Einatomiges ideales Gas. Ein Gas wird als ideal bezeichnet, wenn es so verdünnt ist, daß die Wechselwirkungen zwischen den Teilchen nur für die Energieübertragung durch Stöße, aber nicht für den Energieinhalt eine Rolle spielen. Es sollen N Atome eines solchen Gases in einem Behälter vom Volumen v und der Temperatur T im Gleichgewicht betrachtet werden. Pro cm³ hat ein solches System z. B. bei $T = 300$ K und einer Atommasse von 100 gmol^{-1} etwa $1{,}4 \cdot 10^{27}$ Quantenzustände, während bei 10^5 Pa Druck in diesem Volumen nur etwa $3 \cdot 10^{19}$ Atome enthalten sind. Jeder Quantenzustand wird daher mit sehr kleiner Wahrscheinlichkeit von einem Atom besetzt, während eine gleichzeitige Doppelbesetzung praktisch nicht vorkommt. Es liegt somit eine Boltzmannverteilung (↑ 2.4.4.7.1.) vor.

6.1.3.3.1. *Zustandssumme.* Wegen der Nichtunterscheidbarkeit und Unabhängigkeit der Atome ist die Zustandssumme des Gases

(5) $Z = 1/N!\, Z_A^N$,

wobei Z_A die Zustandssumme eines einzelnen Atoms bedeutet. Sie läßt sich wiederum aus den Zustandssummen über die translatorischen Freiheitsgrade und über die davon unabhängigen aber unterscheidbaren inneren Freiheitsgrade bilden: $Z_A = Z_{tr}Z_{in}$. Bei den meisten Atomen, insbesondere bei den Edelgasatomen, ist mit der Anregung eines inneren Freiheitsgrads ein Energieaufwand von etwa 100 kJmol^{-1} oder mehr verbunden. Bei Temperaturen, für die RT wesentlich kleiner als 100 kJmol^{-1} ist, (bei Raumtemperatur ist z. B. $RT \approx 2{,}5\ \text{kJmol}^{-1}$) beschränkt sich die Zustandssumme über die inneren Freiheitsgrade auf den Summanden mit der niedrigsten Energie, die man ohne Einfluß auf das Endresultat willkürlich gleich Null setzen kann. Dann wird $Z_{in} = 1$.
Z_{tr} läßt sich seinerseits darstellen als $Z_{tr} = Z_{tr,x}Z_{tr,y}Z_{tr,z}$, weil die Bewegungen in den drei Koordinatenrichtungen x, y, z voneinander unabhängig und unterscheidbar sind. Bewegt sich ein Teilchen der Masse m in einem quaderförmigen Kasten der Länge l_x in x-Richtung mit dem Impuls p_x, so ist seine de-Broglie-Wellenlänge (↑ 1.2.3.) $\lambda_x = h/p_x = 2l_x/n_x$, $n_x = 1, 2 \ldots$ Daraus folgt für die möglichen Zustände der Translationsenergie $E_{tr,x} = p_x^2/(2m) = (1/8m)(h^2n_x^2/l_x^2)$ und für die Zustandssumme $Z_{tr,x} = \sum_{n_x=1}^{\infty} \exp[-h^2n_x^2/(8ml_x^2kT)]$. In einem genügend großen Kasten und bei genügend hoher Temperatur unterscheiden sich aufeinanderfolgende Summanden nur wenig, so daß die Zustandssumme durch ein Integral ersetzt werden kann, das $Z_{tr,x} = \frac{(2\pi mkT)^{1/2}}{h} \cdot l_x$ liefert. Da für die y- und z-Richtung analoge Ausdrücke gelten, ergibt sich die translatorische Zustandssumme zu

$$(6) \quad Z_{tr} = \frac{(2\pi mkT)^{3/2}}{h^3} l_x l_y l_z = \frac{(2\pi mkT)^{3/2}}{h^3} v$$

und für die Zustandssumme von N in einem Volumen v eingeschlossenen Atomen eines idealen Gases

$$(7) \quad Z = \frac{(2\pi mkT)^{3N/2}}{N!\, h^{3N}} v^N$$

6.1.3.3.2. *Thermodynamische Zustandsgrößen.* Aus der Zustandssumme Z folgt nach 6.1.(4) für den Druck $p = kTN/v$, was wegen $Nk = nR$ mit der empirischen Zustandsgleichung $pv = nRT$ (↑ 6.2.2.1.) übereinstimmt. R ist die molare Gaskonstante. Die Berechnung der inneren Energie nach 6.1.(1) liefert $u = (3/2)nRT$. Für große Teilchenzahlen N ergibt sich aus 6.1.(3) in guter Näherung

$$s = nR\left(\frac{5}{2} + \ln \frac{(2\pi m)^{3/2}(kT)^{5/2}}{h^3 p}\right).$$

Diese Gleichung für die Entropie eines einatomigen idealen Gases wurde 1912 von SACKUR und TETRODE abgeleitet. Die Entropie hängt von der Masse m der Atome ab (Tab. 6.1.3-1). Die kleinen Abweichungen sind durch die nicht vollständige Idealität der Gase bedingt.

Tab. 6.1.3-1 Entropien einatomiger Gase bei 298,15 K und $1{,}013 \cdot 10^5$ Pa in $\text{Jmol}^{-1}\text{K}^{-1}$

	He	Ne	Ar	Kr
gemessen	127,1	146,5	154,0	164,1
berechnet	125,9	146,0	154,1	166,3

Hätte man die Atome als unterscheidbar betrachtet und dementsprechend in der Zustandssumme den Faktor $1/N!$ weggelassen, so wären alle molaren Entropien um 445 $\text{J mol}^{-1}\text{ K}^{-1}$ größer und damit vollständig falsch geworden. Die aus der *Sackur-Tetrode-Gleichung* berechnete Entropie strebt gegen minus unendlich, wenn sich die Temperatur T dem Wert 0 K nähert. Dies darf nicht verwundern, da die Gleichung unter der Voraussetzung der Gültigkeit der Boltzmannverteilung abgeleitet wurde, die bei sehr tiefen Temperaturen nicht mehr erfüllt ist. Solange sich ein Gas ideal verhält, ist sie aber gültig und erfüllt auch die aus der phänomenologischen Thermodynamik bekannten Gleichungen $S(p_2, T_1) - S(p_1, T_1) = R \ln(p_1/p_2)$ und $S(p_1, T_2) - S(p_1, T_1) = (5/2)\, R \ln(T_2/T_1)$.

6.1.3.4. Zwei- und mehratomige ideale Gase. Der Ausdruck 6.1.(5) für die Zustandssumme des Gases behält seine Gültigkeit, wenn Z_A durch Z_{mol} für ein Molekül ersetzt wird. Bei der Berechnung von Z_{mol} sind jedoch die zusätzlichen inneren Freiheitsgrade bei Molekülen zu berücksichtigen. Dies sind z. B. Schwingungen der Atomkerne gegeneinander und Rotationen des gesamten Moleküls. Die Translation kann wie bei einatomigen Gasen behandelt werden; man hat lediglich die Molekülmassen anstelle der Atommasse einzusetzen. Die erwähnten Freiheitsgrade können mit guter Näherung als unabhängig betrachtet werden. Die Auswirkung der bei der Rotation auftretenden Zentrifugalkräfte auf die Schwingungen braucht erst bei sehr genauen Berechnungen und hohen Temperaturen mitberücksichtigt zu werden. Die Zustandssumme über die inneren Freiheitsgrade eines Moleküls läßt sich in die Anteile $Z_{in} = Z_{el}Z_{schw}Z_{rot}$ zerlegen (↑ 5.3.).

6.1.3.4.1. *Elektronische Zustandssumme.* Bei den meisten stabilen Molekülen liegen sämtliche Elektronenspins gepaart vor. Man findet dann meist den nächsthöheren elektronischen Quantenzustand in einem Abstand von einigen 10^{-19} J. Bei Temperaturen, bei denen das Molekül noch stabil ist, wird es daher nicht wesentlich angeregt. Man erhält $Z_{el} = 1$, wenn der Nullpunkt der Energie in den tiefsten Elektronenzustand verlegt wird.

Bei Radikalen hingegen ist der Grundzustand entsprechend den zwei möglichen Spinzuständen zweifach entartet, und es wird $Z_{el} = 2$. Einen Sonderfall stellt das NO-Molekül dar, in dem im Abstand $\Delta E = 2{,}38 \cdot 10^{-7}$ J oberhalb des tiefsten Elektronenzustands ein zweiter, ebenfalls entarteter Elektronenzustand liegt. Da bei Raumtemperatur $kT \approx 4 \cdot 10^{7}$ J gilt, ist dieser Zustand häufig besetzt, und man erhält als elektronische Zustandssumme $Z_{el,NO} = 2[1 + \exp(-\Delta E/kT)]$. Bei dreifacher Entartung des Grundzustands, wie z. B. beim O_2, wird entsprechend $Z_{el} = 3$.

6.1.3.4.2. *Schwingungszustandssumme.* Ein *zweiatomiges Molekül* kann in guter Näherung als harmonischer Oszillator mit der Frequenz f angesehen werden (↑ 5.3.1.2.). Für seine Energiewerte folgt aus der Quantenmechanik $E_n = [n + (1/2)]\, hf$ mit $n = 0, 1, 2 \ldots$ Wählt man als Nullpunkt der Energie die Energie des tiefsten Zustands, dann ist die Schwingungszustandssumme

$$(8) \quad Z_{schw} = \frac{\exp[-hf/(2kT)]}{1 - \exp[-hf/(kT)]}\,.$$

Wenn $hf \ll kT$ ist, wird $Z_{schw} = kT/(hf)$; wenn $hf \gg kT$ ist, wird $Z_{schw} = \exp[-hf/(2kT)]$, hat also den Wert, der sich ergeben würde, wenn nur das niedrigste Schwingungsenergieniveau mit $n = 0$ existierte.

Z_{schw} nimmt für ein Molekül mit einer bestimmten Schwingungsfrequenz f bei der *charakteristischen Temperatur der Schwingung* $\Theta_{schw} = hf/k$ einen mittleren Wert an. Bei dieser Temperatur ist die Wahrscheinlichkeit, das Molekül im n-ten Schwingungsquantenzustand zu finden, gerade e-mal größer als die Wahrscheinlichkeit, es im Quantenzustand $(n + 1)$ zu finden. In Tab. 6.1.3-2 sind aus spektroskopisch ermittelten Schwingungsfrequenzen berechnete Z_{schw} zusammengestellt. Man sieht, daß bei Raumtemperatur für Moleküle mit schwereren Atomen und dementsprechend niedriger Schwingungsfrequenz Z_{schw} bedeutend werden kann.

Tab. 6.1.3-2 Schwingungszustandssumme für einige Moleküle bei 298,15 K

Molekül	H_2	Cl_2	Rb_2
f/s^{-1}	$13{,}2 \cdot 10^{13}$	$1{,}69 \cdot 10^{13}$	$0{,}171 \cdot 10^{13}$
θ_{schw}/K	5995	801	82
Z_{schw}	$4{,}3 \cdot 10^{-5}$	0,28	3,62

Bei einem *n-atomigen Molekül* hat jede der unabhängigen *Normalschwingungen* (↑ 5.3.1.2.) im allgemeinen ihre eigene Frequenz f_i. Die Schwingungszustandssumme wird dann $Z_{schw} = \Pi\, Z_{schw,i}$. $Z_{schw,i}$ ist die Zustandssumme über die i-te Normalschwingung; sie berechnet sich aus 6.1.(8).

6.1.3.4.3. *Rotationszustandssumme.* Aus der Quantenmechanik (↑ 5.3.1.1.) folgt, daß bei der Rotation eines *zweiatomigen Moleküls* die Energiezustände $E_l = h^2 l(l + 1)/(8\pi^2 J)$ mit $l = 0, 1, 2 \ldots$ möglich sind. J ist das Trägheitsmoment. Die Energiezustände sind entartet. Zu E_l gehören $(2l + 1)$ Quantenzustände. Bei den meisten gasförmigen Stoffen sind ebenso wie bei der Translation sehr viele Rotationszustände besetzt. Die *charakteristische Temperatur der Rotation* $\Theta_{rot} = h^2/(8\pi^2 Jk)$ liegt daher sehr niedrig. Sind die beiden Atome des Moleküls verschieden, so ist die Zustandssumme $Z_{rot} = \sum_{l=0}^{\infty} (2l + 1) \exp[(-\Theta_{rot}/T)\, l(l + 1)]$. Bei genügend hohen Temperaturen $T \gg \Theta_{rot}$ läßt sich diese Summation durch eine Integration ersetzen, und man erhält $Z_{rot} = T/\Theta_{rot} = 8\pi^2 JkT/h^2$. Bei tiefen Temperaturen, vergleichbar mit Θ_{rot}, wird die Rotationszustandssumme größer als die so berechnete. Bei $T = 0$ wird sie gleich Eins. Bei $T = 10\Theta_{rot}$ ist die Gleichung bereits auf 3 % genau.

Bei Molekülen aus zwei gleichen Atomen ist zu beachten, daß zwei um 180° gegeneinander verdrehte Orientierungen nicht unterscheidbar sind. Die Zustandssumme muß dann noch durch die *Symmetriezahl* σ dividiert werden:

$$(9) \quad Z_{rot} = 1/\sigma \cdot T/\Theta_{rot}.$$

Diese gibt die Anzahl der nicht unterscheidbaren Stellungen an, in die ein Molekül durch starre Rotation gebracht werden kann. Für ein aus zwei gleichen Atomen bestehendes Molekül ist somit $\sigma = 2$ und $Z_{rot} = (1/2)\, T/\Theta_{rot}$.

Das Rotationsverhalten von *linearen mehratomigen Molekülen* ist genau gleich dem von zweiatomigen. Die Ausdrücke für die Energiezustände *nichtlinearer Rotatoren* sind relativ kompliziert. Die Zustandssumme wird im Fall $T \gg \Theta_{rot}$ aber geschlossen darstellbar: $Z_{rot} = \frac{16\pi^3}{\sigma h^3} (2\pi J_x J_y J_z)^{1/2} (kT)^{3/2}$. J_x, J_y und J_z sind die Hauptträgheitsmomente (↑ 5.3.1.1.). Die Symmetriezahl σ von H_2O ist 2, von NH_3 3, von CH_4 13, von Benzen 12 und von Naphthalen 4.

6.1.3.4.4. *Zustandssumme über die Kernspins.* Viele Atomkerne haben ein Eigendrehmoment, das sich jedoch bei chemischen Reaktionen

nicht ändert und deshalb meist vernachlässigt werden kann. Die entsprechende Zustandssumme entspricht dem Entartungsgrad der Kernrotationszustände. Für einen Kern mit der Spinquantenzahl s_i ist $Z_{rot,K,i} = 2s_i + 1$ und für ein Molekül mit mehreren unabhängigen Kernen $Z_{rot,K} = \Pi(2s_i + 1)$. $Z_{rot,K}$ fällt bei der Betrachtung chemischer Gleichgewichte stets weg.

6.1.3.4.5. *Chemische Konstitution und Entropie idealer Gase.* Die molare Entropie $S = R(\ln Z + T \cdot \partial \ln Z/\partial T)$ setzt sich ebenso wie die innere Energie U aus Beiträgen der verschiedenen Freiheitsgrade zusammen: $S = S_{tr} + S_{rot} + \sum_i S_{schw,i}$. Der *Beitrag der Translation* wird auch für Moleküle durch die Sackur-Tetrode-Gleichung (↑ 6.1.3.3.2.) beschrieben. Sie zeigt, daß bei gegebener Temperatur und gegebenem Druck der translatorische Entropieanteil für schwerere Moleküle größer ist als für leichte.

Für den *Beitrag der Rotation linearer Moleküle* ergibt sich $S_{rot} = R[\ln T(\sigma\Theta_{rot}) + 1]$. Dieser Anteil ist bei gegebener Temperatur um so größer, je kleiner Θ_{rot}, d. h. je größer das Trägheitsmoment des Moleküls ist. Für ein symmetrisches Molekül ist er bei gleichem Trägheitsmoment um $R \ln 2$ kleiner als für ein nichtsymmetrisches. Für *nichtlineare Moleküle* wird

$$S_{rot} = R\left(\ln \frac{16\pi^3(2\pi J_x J_y J_z)^{1/2}(kT)^{3/2}}{h^3} + \frac{3}{2}\right).$$

Auch hier führt ein großes Produkt der Hauptträgheitsmomente und eine kleine Symmetriezahl σ zu großen Beiträgen zur Entropie.

Für den *Beitrag einer Schwingung i* erhält man

$$(10) \quad S_{schw,i} = R\left[\Theta_{schw,i}/T \cdot \frac{e^{-\Theta_{schw,i}/T}}{1 - e^{-\Theta_{schw,i}/T}} - \ln\left(1 - e^{-\Theta_{schw,i}/T}\right)\right].$$

Diese Funktion zeigt, daß Schwingungen nur dann einen wesentlichen Beitrag zur Entropie liefern, wenn $\Theta_{schw,i}/T < 4$, d. h. wenn bei der Temperatur T die Schwingungsfrequenz $f \lesssim 4\,kT/h$ ist. Bei Raumtemperatur weisen schon kleine Moleküle, wie Cl_2, SO_2, CO_2, solche Schwingungen auf. Größere Moleküle haben in der Regel viele Normalschwingungen, die wesentlich zur Entropie beitragen.

Je kleiner die Frequenz einer Normalschwingung ist, um so größer wird bei gegebener Temperatur ihr Beitrag zur Entropie. Für $f_i \to 0$ würde er nach 6.1.(10) gegen Unendlich streben. In solchen Fällen sind jedoch sehr viele Schwingungszustände angeregt; die Näherung durch einen harmonischen Oszillator wird schlecht, und damit verliert diese Gleichung ihre Gültigkeit. Wichtig sind Fälle, in denen bestimmte Normalschwingungen in mehr oder weniger freie innere Rotationen ausarten. Man hat dann eine Zustandssumme über die Zustände dieser inneren Rotation zu bilden. Der sich daraus ergebende Beitrag zur Entropie kann beträchtlich sein.

Im allgemeinen ist die Entropie um so größer, je größer die Masse und das Trägheitsmoment sind und je weniger starr und symmetrisch ein Molekül gebaut ist (Tab. 6.1.3-3).

Tab 6.1.3-3 Molare Standardentropie $S^\ominus$ einiger Gase bei 298 K in $Jmol^{-1}K^{-1}$

	$S^\ominus$
$H_3C—CH_2—CH_3$	270
Cyclopropan: H_2C, CH_2 und $C H_2$ im Dreiring	239
$C(CH_3)_4$ (vier CH_3 an C)	307
$H_3C—CH(CH_3)—CH_2—CH_3$	343
$H_3C—CH_2—CH_2—CH_2—CH_3$	350

6.1.3.4.6. *Homogene Gasreaktionen im Gleichgewicht.* Für die Gleichgewichtskonstante K einer homogenen Reaktion zwischen idealen Gasen gilt $\Delta_R G^\ominus = -RT \ln K$ (↑ 6.1.2.4.3.). Man hat somit die freie Standardreaktionsenthalpie $\Delta_R G^\ominus$ zu berechnen. Für eine Reaktion $\nu_A A + \nu_B B \rightarrow \nu_C C + \nu_D D$ ist $\Delta_R G^\ominus = \nu_C G_C^\ominus + \nu_D G_D^\ominus - \nu_A G_A^\ominus - \nu_B G_B^\ominus$. Mit $G_i = RT \ln N_A - RT \ln Z_i$ erhält man daraus

$$K = \left(\frac{1}{N_A}\right)^{\nu_C + \nu_D - \nu_A - \nu_B} \frac{Z_C^{\nu_C} \cdot Z_D^{\nu_D}}{Z_A^{\nu_A} \cdot Z_B^{\nu_B}}.$$

N_A ist die Avogadrokonstante.

Die zur Berechnung der molaren Zustandssummen benötigten Elektronenenergien, Trägheitsmomente und Schwingungsfrequenzen können in vielen Fällen aus Absorptionsspektren experimentell bestimmt werden (↑ 5.3.). Die rein theoretische Berechnung dieser Größen mit Hilfe der Quantentheorie ist gegenwärtig nur für sehr einfache Moleküle, z. B. H_2, mit solcher Genauigkeit möglich, daß sich daraus zuverlässige Werte für Gleichgewichtskonstanten ableiten lassen. Die Gleichung bietet deshalb vor allem die Möglichkeit, aus spektroskopischen Daten Gleichgewichtskonstanten für Temperaturen zu berechnen, bei denen sie der direkten Messung nicht zugängig sind.

Zum Beispiel erhält man für das Dissoziationsgleichgewicht des Kaliums $K_2 \rightleftharpoons 2\,K$ bei 973,2 K

$$(11) \quad K = \left(\frac{1}{N_A} \frac{Z_{tr,K}^2}{Z_{tr,K_2}}\right) \frac{1}{Z_{schw,K_2}} \frac{1}{Z_{rot,K_2}} \frac{Z_{el,K}^2}{Z_{el,K_2}}.$$

Die Massen des K-Atoms und des K_2-Moleküls betragen $6{,}25 \cdot 10^{-23}$ g bzw. $12{,}5 \cdot 10^{-23}$ g. Damit erhält man für den ersten Faktor beim Standarddruck $p^{\ominus} = 1{,}013 \cdot 10^5$ Pa (= 1 atm) nach 6.1.(6) $Z^2_{\text{tr,K}}/(N_A Z_{\text{tr,K}_2}) = 0{,}68 \cdot 10^8$.
Aus der Schwingungsfrequenz des K_2-Moleküls $f = 2{,}779 \cdot 10^{12}\ \text{s}^{-1}$ berechnet sich der zweite Faktor [↑ 6.1.(8)] zu $1/Z_{\text{schw,K}_2} = 1/7{,}37$. Das Trägheitsmoment des K_2-Moleküls $J = 4{,}92 \cdot 10^{-38}$ gcm² wurde aus spektroskopischen Messungen bestimmt. Da das Molekül aus zwei gleichen Atomen besteht, ist seine Symmetriezahl $\sigma = 2$. Daraus folgt für den dritten Faktor [↑ 6.1.(9)] $1/Z_{\text{rot,K}_2} = 1/(5{,}94 \cdot 10^3)$.
Für die elektronische Dissoziationsenergie des K_2-Moleküls ergibt sich aus spektroskopischen Beobachtungen $D_{\text{el}} = 0{,}833 \cdot 10^{-19}$ J. Das K_2-Molekül hat nur gepaarte Elektronenspins. Seine elektronische Zustandssumme ist daher $Z_{\text{el,K}_2} = \exp(D_{\text{el}}/kT)$.
Das K-Atom liegt in einem Dublettzustand vor, so daß seine Zustandssumme bei dem gewählten Energienullpunkt $Z_{\text{el,K}} = 2$ wird. Der vierte Faktor in 6.1.(11) ergibt sich daher zu $Z^2_{\text{el,K}}/Z_{\text{el,K}_2} = 2^2/(4{,}93 \cdot 10^2)$. Das Produkt aller vier Faktoren ergibt schließlich für die Gleichgewichtskonstante $K = 12{,}5$.
Bei 973,2 K beträgt der Dampfdruck von Kalium $p = 0{,}564 \cdot 10^5$ Pa. Daraus kann man den Dissoziationsgrad α des gesättigten Dampfs berechnen. Aus $K = 4\alpha^2 p/(1 - \alpha^2) = 12{,}5$ ergibt sich $\alpha = 0{,}922$, d. h. 4,1 % der Teilchen im gesättigten Kaliumdampf sind K_2-Moleküle. Dieser aus spektroskopischen Daten bestimmte Wert von α stimmt gut mit dem in diesem Fall direkt beobachteten Wert $\alpha_{\text{exp}} = 0{,}920$ überein.
Ein anderes Beispiel ist der Einfluß von Isotopen auf das Gleichgewicht. Es seien die zwei Dissoziationsreaktionen $RH \rightleftharpoons R^- + H^+$ und $RD \rightleftharpoons R^- + D^+$ mit den Gleichgewichtskonstanten K_H bzw. K_D betrachtet. RH sei ein größeres Säuremolekül, RD das entsprechende deuterierte Molekül. Die Änderung der Elektronenenergie ist bei beiden Reaktionen die gleiche. Trägheitsmomente und Massen von RH und RD unterscheiden sich nur wenig. Der Hauptunterschied besteht in den Schwingungsfrequenzen des Wasserstoffatoms und des Deuteriumatoms gegenüber dem Rest R.
Das Verhältnis der Dissoziationskonstanten ist $K_H/K_D = (Z_{H^+}/Z_{D^+})(Z_{RD}/Z_{RH})$. Der erste Faktor wird entsprechend 6.1.(6) $Z_{H^+}/Z_{D^+} = (m_{H^+}/m_{D^+})^{3/2} = 1/2^{3/2}$. Im zweiten Faktor können mit guter Näherung $Z_{\text{tr,RH}} = Z_{\text{tr,RD}}$ und $Z_{\text{rot,RH}} = Z_{\text{rot,RD}}$ gesetzt werden. Damit wird $Z_{RD}/Z_{RH} \approx Z_{\text{schw,RD}}/Z_{\text{schw,RH}} \approx \exp[-h \sum_i (f_{i,\text{RD}} - f_{i,\text{RH}})/(2kT)]$. Dieser Zusammenhang gilt, solange die charakteristischen Temperaturen der durch die Isotopen bestimmten Schwingungen wesentlich oberhalb der Versuchstemperaturen liegen. Wird an einem Kohlenstoffatom oder an einem Sauerstoffatom eines organischen Stoffs ein H-Atom durch ein D-Atom ersetzt, so ändern sich hauptsächlich die Frequenzen der Streckschwingungen und der zwei in zueinander senkrechten Ebenen erfolgenden Biegeschwingungen. Mittelwerte aus spektroskopischen Beobachtungen liefern schließlich für das Verhältnis der Dissoziationskonstanten bei Raumtemperatur $K_H/K_D \approx 27$. Die größere Konstante von RH ist im wesentlichen auf die größere Nullpunktsenergie der R-H-Schwingung zurückzuführen. Der niedrigste Energiezustand der langsameren R-D-Schwingung liegt tiefer. Für die Dissoziation von RD wird deshalb mehr Energie benötigt. Diese Abschätzungen gelten jedoch nur für den Gaszustand. In flüssiger Phase ist der Isotopeneffekt abgeschwächt, weil die D^+-Ionen mit kleinerer Nullpunktsschwingungsenergie an die Lösungsmittelmoleküle angelagert werden als die Protonen. Dann ist gewöhnlich $K_H/K_D \approx 8$.

6.1.3.5. Reale Gase. Wenn die Teilchendichte in einem Gas größer wird, treten infolge der Kraftwirkungen zwischen den Teilchen Abweichungen von der thermischen Zustandsgleichung des idealen Gaszustands (↑ 6.2.2.1.) auf. Bisher ist es jedoch noch nicht gelungen, die Abhängigkeit dieser Kräfte vom Abstand und der gegenseitigen Orientierung der Teilchen aus den Zustandsgleichungen realer Gase direkt zu bestimmen. Man versucht deshalb umgekehrt, plausible Modellvorstellungen über die Abstandsabhängigkeit der Wechselwirkungsenergien zu entwickeln und in Formeln darzustellen (↑ 6.2.4.1.). Diese enthalten bestimmte Parameter, deren Werte sich aus den Zustandsgleichungen realer Gase bestimmen lassen. Die statistische Thermodynamik ermöglicht es dann, Zusammenhänge zwischen den interkorpuskularen Kräften und den makroskopischen Eigenschaften abzuleiten.
Bezogen auf ein fest mit den Wänden eines Gasbehälters verbundenes Koordinatensystem ist die mittlere Teilchendichte eines Gases $\bar{N} = N_A/V$ überall gleich. N_A ist die Avogadrokonstante und V das Molvolumen des Gases. In einem Koordinatensystem, dessen Ursprung sich mit einem Gasteilchen bewegt, ist die Teilchendichte N jedoch ortsabhängig. Abgesehen von der Anisotropie der Teilchen hängt sie vom radialen Abstand r vom Bezugsteilchen ab. In sehr kleinen Abständen, wo das Abstoßungspotential sehr groß ist, können sich keine weiteren Teilchen aufhalten. In Bereichen der Anziehung, z. B. in der Umgebung des Minimums des Lennard-Jones-Potentials (↑ 6.2.4.1.), ist die Teilchendichte größer als $\bar{N}$. Bei sehr großen Abständen wird sie gleich $\bar{N}$.
Die Funktion $N(r)$ wird *radiale Dichteverteilung*

genannt. Sie kann in guter Näherung durch $N(r) = \bar{N} \cdot \exp[-\varphi(r)/(kT)]$ dargestellt werden, wenn die Temperatur genügend hoch und die Teilchendichte im Gas noch so klein ist, daß sich im Kraftfeld des Bezugsteilchens nur sehr selten zwei andere Teilchen befinden, die sich ihrerseits wieder beeinflussen. $\varphi(r)$ ist die potentielle Wechselwirkungsenergie zwischen zwei Teilchen. Für ein Lennard-Jones-Potential nimmt die radiale Dichteverteilung die in Abb. 6.1.3-4 dargestellte Form an.

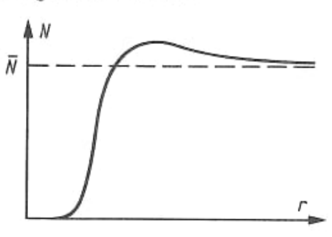


Abb. 6.1.3-4 Qualitativer Verlauf der radialen Dichteverteilung für ein Lennard-Jones-Potential

Die interkorpuskularen Wechselwirkungen führen dazu, daß die molare innere Energie U und daraus abgeleitete thermodynamische Zustandsgrößen von der mittleren Teilchendichte $\bar{N}$ und damit vom Molvolumen V abhängig werden.

6.1.3.6. Flüssigkeiten. In Flüssigkeiten besteht eine *Nahordnung* der Teilchen, jedoch keine Fernordnung. Benachbarte Teilchen sind nicht unabhängig voneinander. Die Zustandssumme läßt sich deshalb nicht mehr als Produkt molekularer Zustandssummen darstellen. Ihre exakte Berechnung ist bis heute noch nicht gelungen. Man muß daher das makroskopische Verhalten aufgrund vereinfachender Modellvorstellungen zu verstehen versuchen. Dies gelingt insbesondere durch Näherungsansätze zur Berechnung der radialen Dichteverteilung (↑ 6.1.3.5.) oder der Zustandssumme sowie deren Druck- und Temperaturabhängigkeiten.

Da die interkorpuskularen Kräfte im allgemeinen kurze Reichweiten haben, ist für die Berechnung der inneren Energie die kristallähnliche nähere Umgebung jedes Teilchens besonders wichtig. Ein von der Kristallstruktur ausgehendes anschauliches und realistisches Modell ist das Löchermodell.

6.1.3.6.1. *Löchermodell der Flüssigkeiten.* Aus experimentell bestimmbaren radialen Dichteverteilungen in Flüssigkeiten geht hervor, daß die mittlere Zahl der nächsten Nachbarn eines Teilchens mit steigender Temperatur abnimmt. Man kann deshalb annehmen, daß bei kristallähnlicher Nahordnung in der Nachbarschaft einzelner Teilchen ein Gitterplatz unbesetzt ist. Die Bildung eines solchen Lochs erfordert eine Arbeit w. Sie setzt sich zusammen aus einem Anteil w_0, der benötigt wird, um ein Teilchen aus dem Kristallverband zu entfernen und an der Oberfläche anzulagern, und einem Anteil pV_0/N_A für die Volumenvergrößerung V_0/N_A beim Außendruck p: $w = w_0 + pV_0/N_A$. Unter der Voraussetzung, daß die Zahl der Löcher wesentlich kleiner ist als die Zahl der besetzten Gitterplätze, ergibt sich für die Wahrscheinlichkeit, daß ein bestimmter Gitterplatz unbesetzt ist $N_L/(N_A + N_L) = \exp[-w/(kT)]/(1 + \exp[-w/(kT)])$. Dabei ist N_L die Zahl der Löcher, N_A die Avogadrokonstante und $1 + \exp[-w/(kT)]$ die Zustandssumme für einen Gitterplatz, der die beiden Zustände „besetzt" mit der Energie Null und „unbesetzt" mit der Energie w annehmen kann. Daraus folgt $N_L = N_A \cdot \exp[-w/(kT)]$ und für das Molvolumen V der Flüssigkeit $V = V_0 + \frac{V_0}{N_A} N_L = V_0\left(1 + \exp\left[\frac{-w_0 - V_0 p/N_A}{kT}\right]\right)$.

Die meisten Stoffe zeigen beim Schmelzen eine Volumenzunahme von etwa 10%, d. h. $\exp[-w/(kT)] \approx 0{,}1$. Bei einer Schmelztemperatur von einigen hundert K und für nicht zu hohe Drücke ergibt sich daraus $w \approx w_0 \approx 10^{-20}$ J. Andererseits ist die Wechselwirkungsenergie zwischen zwei Teilchen in der Flüssigkeit von der Größenordnung 10^{-21} J, so daß bei Berücksichtigung der Zahl der Nachbarn eines Teilchens der empirisch gefundene Wert w_0 physikalisch sinnvoll ist.

Aus der thermischen Zustandsgleichung von Flüssigkeiten läßt sich der isotherme Kompressibilitätskoeffizient χ (↑ 6.1.2.1.4.) abschätzen. Bei einem Molvolumen von 50 $cm^3 mol^{-1}$ und Raumtemperatur wird $\chi \approx 2 \cdot 10^{-9}$ Pa^{-1}. Die Übereinstimmung mit experimentell ermittelten Kompressibilitätskoeffizienten berechtigt zu der Annahme, daß bei der Kompression einer Flüssigkeit vor allem die mittlere Zahl der Nachbarn eines Teilchens erhöht wird, während der mittlere Abstand der Teilchen nicht wesentlich verkleinert wird.

Die Abschätzung des isobaren thermischen Ausdehnungskoeffizienten (↑ 6.1.2.1.4.) ergibt bei Raumtemperatur $\beta \approx 10^{-3}$ K^{-1}. Dieser Wert stimmt in der Größenordnung mit gemessenen Werten überein. Die thermische Ausdehnung von Flüssigkeiten ist demnach vor allem durch die Vermehrung der Zahl der Löcher und nicht durch die Vergrößerung der Abstände zwischen Nachbarteilchen bedingt.

6.1.3.6.2. *Mischungen von Flüssigkeiten.* Liegen in der Mischung nur neutrale Teilchen vor, so braucht nur die Wechselwirkungsenergie zwischen nächsten Nachbarn berücksichtigt zu werden, weil die Kräfte zwischen den Teilchen mit steigendem Abstand voneinander rasch abnehmen. 1 Mol einer binären Mischung besteht

aus $x_1 N_A$ Teilchen der Komponente 1 und $(1 - x_1) N_A$ Teilchen der Komponente 2. Dabei ist x_1 der Molenbruch der Komponente 1 und N_A die Avogadrokonstante. Jedes Teilchen ist im Mittel von m nächsten Nachbarn umgeben, von denen im Mittel mx_1 der Komponente 1 und $m(1 - x_1)$ der Komponente 2 angehören. Die Wechselwirkungsenergie von zwei Teilchen der Komponente 1 sei E_{11}, diejenige von zwei Teilchen der Komponente 2 sei E_{22} und die von zwei verschiedenen Teilchen sei E_{12}.

Der Anteil der interkorpuskularen Wechselwirkung an der molaren inneren Energie ist dann

$U'_M = (1/2)N_A m x_1^2 E_{11} + (1/2)N_A m(1 - x_1)^2 E_{22} + N_A m x_1 (1 - x_1) E_{12}$.

Bei getrennten reinen Komponenten gilt

$U'_{\text{rein}} = (1/2)N_A m x_1 E_{11} + (1/2)N_A m(1-x_1)E_{22}$.

Daraus folgt

$\Delta_M U' = U_M - U'_{\text{rein}} = (1/2)N_A m x_1(1 - x_1)(2E_{12} - E_{11} - E_{22})$.

Bei Raumtemperatur sind außer den elektronischen Freiheitsgraden und den Schwingungsfreiheitsgraden alle Freiheitsgrade voll angeregt und liefern somit einen von der Struktur unabhängigen Beitrag zur molaren inneren Energie. Außerdem ist aus IR-Spektren zu entnehmen, daß die Schwingungsenergien in der Mischung nicht wesentlich von denen der reinen Komponenten abweichen. Der Beitrag der in der Mischung mitunter beträchtlich veränderten elektronisch angeregten Zustände ist wegen ihrer hohen Energieniveaus gering. Man kann somit $\Delta_M U'$ als *innere Mischungsenergie* $\Delta_M U$ betrachten. Die *Mischungsenthalpie* $\Delta_M H = \Delta_M U + p\Delta_M V$ unterscheidet sich bei flüssigen Mischungen nur wenig von $\Delta_M U$, da bei nicht zu hohen Drücken der Term $p\Delta_M V$ vernachlässigbar wird.

Die *Mischungsentropie* $\Delta_M S$ läßt sich wie folgt abschätzen: Wenn man von der relativ kleinen Anzahl von Löchern absieht, stehen in der Mischung N_A Plätze für die Teilchen zur Verfügung. Auf diesen Plätzen gibt es $W = \frac{N_A!}{(x_1 N_A)!\,[N_A(1 - x_1)]!}$ unterscheidbare Anordnungen der $x_1 N_A$ Teilchen der Komponente 1 und der $(1 - x_1)N_A$ Teilchen der Komponente 2. Da die meisten dieser Anordnungen nahezu die gleiche Energie haben, kann W als die thermodynamische Wahrscheinlichkeit des Systems betrachtet werden. Sie ist in den reinen Komponenten gleich Eins. Beim absoluten Nullpunkt würde somit $\Delta_M S = k(\ln W - 0) = -R[x_1 \ln x_1 + (1 - x_1) \ln (1 - x_1)]$.

Dies entspricht auch der Mischungsentropie bei endlichen Temperaturen, wenn die Mischung ideal ist (↑ 6.3.3.3.) oder näherungsweise, wenn die Abweichung von der Idealität nicht allzu stark ist.

6.2. Gase

6.2.1. Charakteristische Merkmale und Einteilung

Im gasförmigen Aggregatzustand sind die Atome oder Moleküle im Mittel sehr weit gegenüber ihren Abmessungen voneinander entfernt und in völlig regelloser Translationsbewegung. Durch Stöße untereinander und auf die Gefäßwände verändern sich ständig ihre Bewegungsrichtungen und Geschwindigkeiten. Gase füllen daher jeden zur Verfügung stehenden Raum vollständig und gleichmäßig aus; sie sind unbegrenzt formveränderlich und miteinander mischbar.

Das makroskopische Verhalten der Gase wird weitgehend durch die *thermische Molekularbewegung* bestimmt. Die *zwischenmolekularen Wechselwirkungskräfte* (↑ 5.1.4.) spielen dagegen im allgemeinen eine wesentlich geringere Rolle. Man kann sich einen Grenzzustand vorstellen, in dem die Teilchen überhaupt keine Kräfte mehr aufeinander ausüben. Dieser *ideale Gaszustand* läßt sich um so besser annähern, je geringer die Gasdichte bzw. der Gasdruck und je höher die Temperatur ist. Dann ist auch das Eigenvolumen aller Teilchen gegenüber dem Volumen des zur Verfügung stehenden Raums vernachlässigbar.

Alle tatsächlich vorkommenden *realen Gase* weisen mehr oder weniger große Abweichungen vom Zustand des idealen Gases auf; die stoffeigenen Wechselwirkungskräfte und Eigenvolumina der Teilchen sind dann nicht mehr zu vernachlässigen.

Mit zunehmendem Druck und abnehmender Temperatur nähern sich alle Gase Zuständen, in denen sowohl die Molekularbewegung als auch die Packungsdichte und damit die Wechselwirkungskräfte den Verhältnissen in Flüssigkeiten vergleichbar werden. Der Übergang in den flüssigen Aggregatzustand erfolgt bei den einzelnen Gasen bei unterschiedlichen Zustandsbedingungen. Besonders hohe Drücke und niedrige Temperaturen sind zur Verflüssigung der *permanenten Gase* H_2, N_2, O_2 und der Edelgase erforderlich. Im normalen Druck- und Temperaturbereich verhalten sich daher diese Gase bereits weitgehend ideal.

Gase bei oder nahe bei ihrem *Kondensations-* oder *Verflüssigungspunkt* bezeichnet man häufig als Dämpfe; im täglichen Sprachgebrauch versteht man darunter auch den Gaszustand von Stoffen, die bei normalem Druck flüssig oder fest vorliegen. Ein Dampf ist *gesättigt*, wenn er mit seiner eigenen kondensierten Phase im thermodynamischen Gleichgewicht steht. Liegt er in einer höheren Konzentration vor, nennt man ihn *übersättigt*. Er ist dagegen *ungesättigt*, wenn weitere flüssige oder feste Mengen des gleichen Stoffs ebenfalls in den Gaszustand übergehen.

6.2.2. Thermische Zustandseigenschaften der Gase

6.2.2.1. Gesetze des idealen Gaszustands. Der *thermische Zustand* reiner Gase läßt sich vollständig durch die Zustandsgrößen Volumen v, Druck p, Temperatur T und Masse m beschreiben. Die Abhängigkeiten dieser Größen voneinander wurden auf empirischem Wege gefunden und als *Grenzgesetze für den idealen Gaszustand* erkannt.

6.2.2.1.1. *Boyle-Mariottesches Gesetz der Isotherme.* Bei konstanter Temperatur ist das Produkt aus Druck und Volumen für eine bestimmte Masse eines Gases eine Konstante, $(pv)_{T,m} = \text{const.}$ Bei 0 °C verhalten sich viele Gase unter niedrigem Druck praktisch diesem Gesetz entsprechend (Abb. 6.2.2-1).

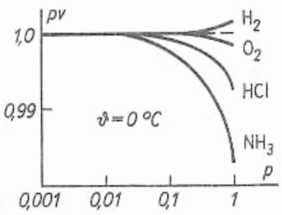


Abb. 6.2.2-1 pv-p-Diagramm von verschiedenen Gasen; Druck p in 10^5 Pa, Volumen v in l, Ordinateneinteilung willkürlich

6.2.2.1.2. *Gay-Lussacsches Gesetz der Isobare.* Eine Gasmenge von konstantem Druck dehnt sich bei einer Temperatursteigerung um ein Kelvin stets um 1/273,15 ihres Volumens v_0 bei $T_0 = 273$ K aus, $v_{p,m} = v_0(1 + \gamma\vartheta) = v_0 T/T_0$. Mit Hilfe des *thermischen Ausdehnungskoeffizienten* γ wird die *absolute Temperatur* $T = 273{,}15\ \text{K} + \vartheta$ definiert, die dieselbe Einheit Kelvin hat wie die Celsiustemperatur ϑ, deren Nullpunkt aber bei $-273{,}15$ °C liegt. Damit

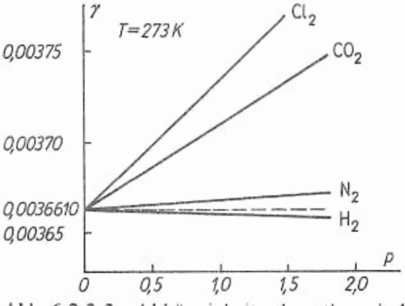


Abb. 6.2.2-2 Abhängigkeit des thermischen Ausdehnungskoeffizienten γ in K^{-1} verschiedener Gase vom Druck p in 10^5 Pa

besagt das *Gay-Lussacsche Gesetz*: Bei konstantem Druck ist das Volumen einer gegebenen Gasmenge der absoluten Temperatur direkt proportional. Tatsächlich nähert sich γ bei hinreichend kleinen, von Gas zu Gas unterschiedlichen Drücken dem „idealen" Wert $1/273{,}15\ K^{-1} = 0{,}003\,661\,0\ K^{-1}$ (Abb. 6.2.2-2).

6.2.2.1.3. *Gay-Lussacsches Gesetz der Isochore.* Wegen des Boyle-Mariotteschen Gesetzes gilt analog, daß eine Gasmenge von konstantem Volumen bei einer Temperatursteigerung um ein Kelvin stets die gleiche Druckänderung um 1/273,15 ihres Drucks p_0 bei 0 °C erfährt, $p_{v,m} = p_0(1 + \beta\vartheta) = p_0 T/T_0$. Der *thermische Spannungskoeffizient* β ist demnach bei idealen Gasen mit dem Ausdehnungskoeffizienten γ identisch.

6.2.2.1.4. *Avogadrosches Gesetz.* Unter gleichen Bedingungen von p und T enthalten gleiche Volumina aller idealen Gase die gleiche Anzahl von Teilchen. Dementsprechend nimmt je 1 Mol mit der *molaren Teilchenzahl* $N_A = 6{,}0225 \cdot 10^{23}\ \text{mol}^{-1}$ der verschiedensten Gase bei gleichem Druck und gleicher Temperatur das gleiche Volumen ein. N_A wird *Avogadrokonstante* genannt. Beim Normaldruck $p_0 = 101{,}325 \cdot 10^3$ Pa (1 atm) und der *Normaltemperatur* $T_0 = 273{,}15$ K (0 °C) ist das *molare Volumen* aller idealen Gase $V_0 = 22{,}4136\ \text{lmol}^{-1}$. Das durch den Quotienten aus *molarer Masse M* und Dichte gegebene molare Volumen einer Reihe von Gasen liegt unter diesen Bedingungen schon nahe bei dem „idealen" Wert.

6.2.2.1.5. *Thermische Zustandsgleichung der idealen Gase.* Aus dem Boyle-Mariotteschen, dem Gay-Lussacschen und dem Avogadroschen Gesetz ergibt sich für 1 Mol eines idealen Gases $V_T p_T/T = V_0 p_0/T_0 = R$. V_T ist das molare Volumen des Gases bei der Temperatur T und dem Druck p_T, V_0 das molare Volumen beim Normaldruck p_0 und der Normaltemperatur T_0. Die universelle Konstante R heißt *molare Gaskonstante* und ergibt sich zu $R = 8{,}3143$ $\text{JK}^{-1}\text{mol}^{-1}$. Für eine Stoffmenge n gilt dann mit $v = nV$ die Beziehung $pv = nRT$.

Diese *Gleichung des idealen Gases* ist ein Grenzgesetz, das von realen Gasen um so besser erfüllt wird, je geringer ihr Druck und je höher ihre Temperatur ist, im Bereich um Normalbedingungen oft mit einer für praktische Zwecke hinreichenden Genauigkeit von wenigen Prozenten.

6.2.2.1.6. *Gasgemische.* Bei idealen Gasgemischen verhält sich jede Gasart i so, als ob sie allein im Gesamtvolumen enthalten wäre, und es gilt für sie $p_i v = n_i RT$. Ebenso wie sich die Gesamtstoffmenge n der Mischung additiv aus den Stoffmengen n_i der einzelnen Komponenten zusammensetzt, gilt für den Gesamtdruck p der Mischung $p = p_1 + p_2 + \dots p_i = \sum p_i$, was als *Daltonsches Partialdruckgesetz* bezeichnet

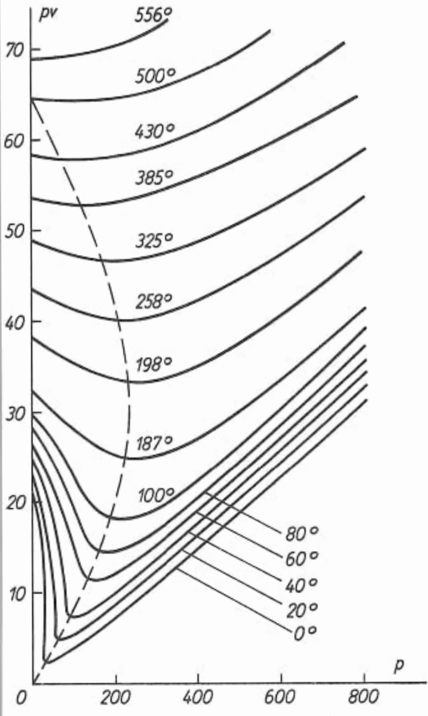


Abb. 6.2.2-3 *pv-p*-Isothermen von CO_2; Temperatur in °C, Druck p in 10^5 Pa, Volumen v in $l\,mol^{-1}$

wird. Die *Partialdrücke* p_i lassen sich aus dem Gesamtdruck nach $p_i = x_i p$ (x_i Molenbruch der Komponente *i*) ermitteln. Für Mischungen realer Gase hat die Gleichung $pv = \sum n_i RT$ wiederum nur Näherungscharakter, wogegen das Partialdruckgesetz für beliebige Gasmischungen gültig ist.

6.2.2.2. Verhalten realer Gase. Schon unter Normalbedingungen weicht das makroskopische Verhalten vieler Gase merklich von dem des idealen Gaszustands ab; z. B. hängt das Produkt *pv* stark vom Druck und von der Temperatur ab (Abb. 6.2.2-3). Wie im Beispiel durchlaufen die Isothermen oft Minima, die mit steigender Temperatur immer flacher werden. Oberhalb der *Boyletemperatur* (für CO_2 500 °C), bei der das Gas bis zu relativ hohen Drücken dem Boyle-Mariotteschen Gesetz gehorcht, steigen die Isothermen monoton an.

Da die Eigenvolumina der Gasteilchen und die zwischenmolekularen Wechselwirkungen von der Natur des Gases abhängen, läßt sich das Verhalten realer Gase nicht mehr durch eine universelle Zustandsgleichung beschreiben. Hierfür sind vielmehr Gleichungen erforderlich, die zusätzliche individuelle Konstanten enthalten.

6.2.2.2.1. *Van-der-Waalssche Zustandsgleichung.* Sie hat zwar heute nur noch beschränkte praktische Bedeutung, läßt aber besonders gut den Einfluß der zwischenmolekularen Wechselwirkungskräfte und der Eigenvolumina der Teilchen erkennen. Dafür wurden von VAN DER WAALS in der Gleichung des idealen Gases zwei Korrekturen an den Größen Druck und Volumen angebracht: $(p + an^2/v^2)\,(v - nb) = nRT$; a und b sind von der Gasart abhängige Stoffwerte, die empirisch ermittelt werden. Das Korrekturglied $an^2/v^2 = a/V^2$ heißt *Binnen-* oder *Kohäsionsdruck* und gibt den Einfluß der Anziehungskräfte zwischen den Gasteilchen an. Diese wirken wie ein nach dem Inneren der Gasphase gerichteter Zug, der sich in einer Verringerung des Gasdrucks gegenüber dem idealen Zustand äußert. Das *Kovolumen b* berücksichtigt das Eigenvolumen der Teilchen; für annähernd kugelförmige Atome oder Moleküle ist b etwa gleich dem Vierfachen des molaren Teilchenvolumens.

Abb. 6.2.2-4 zeigt, daß oberhalb einer bestimmten, von Gas zu Gas unterschiedlichen Temperatur T_k die *p-V*-Isothermen wie bei einem idealen Gas monoton verlaufen, das Volumen also eine eindeutige Funktion des Drucks ist. Alle Isothermen unterhalb dieser Temperatur weisen einen S-förmigen Verlauf auf. Der Kurventeil *A–B* widerspricht aber der Erfahrung, denn das Volumen müßte mit fallendem Druck zunehmen. In Wirklichkeit wird beim Komprimieren eines Gases die Isotherme bei *D* verlassen, der Druck bleibt zwischen *D* und *C* konstant, um danach sehr steil anzusteigen. In Punkt *D* beginnen Flüssigkeitströpfchen aufzutreten, von *D* bis *C* wird das Gas zunehmend verflüssigt, bei *C* liegt nur noch Flüssigkeit vor.

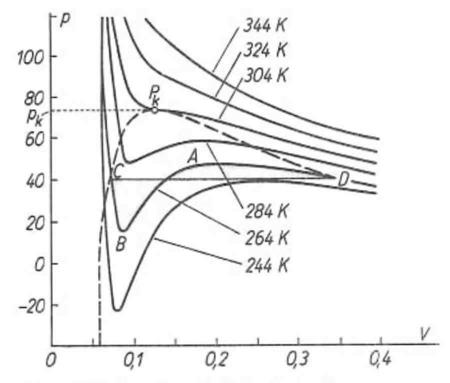


Abb. 6.2.2-4 Van-der-Waals-Isothermen von CO_2; Druck p in 10^5 Pa, Molvolumen V in l, $p_k = 73{,}9 \cdot 10^5$ Pa

Die van-der-Waalssche Zustandsgleichung beschreibt demnach nicht nur das Verhalten von realen Gasen, sondern auch von Flüssigkeiten. Sie versagt allerdings, wenn Gas und Flüssigkeit nebeneinander vorliegen. Mit steigender Temperatur wird dieses Zweiphasengebiet immer kleiner, um bei der *kritischen Temperatur* T_k in den *kritischen Punkt* P_k überzugehen. Oberhalb P_k ist eine Verflüssigung des Gases auch bei noch so hohen Drücken unmöglich. Die stoffeigenen *kritischen Daten* T_k und p_k ermöglichen eine Berechnung der Größen a und b der van-der-Waalsschen Gleichung.

Eine noch bessere Widerspiegelung des realen Gaszustands ist durch Zustandsgleichungen mit einer größeren Anzahl stoffspezifischer Konstanten möglich. Oft benutzt werden z. B. die Gleichungen von BEATTY und BRIDGMAN mit fünf oder von BENEDICT, WEBB und RUBIN mit acht empirischen Konstanten.

6.2.2.2.2. *Virialgleichung*. In dieser Beziehung wird die Abhängigkeit des Produkts pV vom Druck durch eine Potenzreihe in p berücksichtigt: $pV = RT + B(T)\,p + C(T)\,p^2 + \ldots$ Die temperaturabhängigen Stoffgrößen $B(T)$, $C(T)$ usw. heißen *zweiter*, *dritter* usw. *Virialkoeffizient*. Sie können, auch für Gasmischungen, experimentell bestimmt werden. Aus ihnen sind auch Aussagen über die Potentialfunktionen der Wechselwirkungen zwischen den Gasteilchen möglich.

6.2.2.2.3. *Korrespondenzprinzip*. Eine universell gültige Zustandsgleichung für den realen Gaszustand könnte dann zu erwarten sein, wenn die Potentialfunktionen der zwischenmolekularen Wechselwirkungen zwischen den Teilchen möglichst gleiche Gestalt haben und ein einheitlicher Bezugszustand existiert, in dem die verschiedenen Gase vergleichbare Eigenschaften haben. Dies sollte am ehesten an den kritischen Punkten der Fall sein. Die Abweichungen von diesem Bezugszustand sollten dann für alle Gase einheitlich dargestellt werden können, indem an Stelle von p, v und T modifizierte Zustandsgrößen eingeführt werden, die den individuellen Daten für den gemeinsamen Bezugszustand Rechnung tragen. Als sehr zweckmäßig haben sich die *reduzierten*, d. h. durch die entsprechenden kritischen Daten dividierten, *Zustandsgrößen* $\pi = p/p_k$, $\varphi = v/v_k$ und $\tau = T/T_k$ erwiesen. Zwischen diesen sollte nach dem Korrespondenzprinzip oder *Theorem der übereinstimmenden Zustände* ein funktionaler Zusammenhang in Form einer allgemeingültigen thermischen Zustandsgleichung bestehen. Das trifft für Stoffe aus Teilchen ohne stärkere gerichtete Wechselwirkungskräfte tatsächlich auch zu. Durch Hinzunahme eines, bis heute allerdings nur empirisch aus Dampfdruckmessungen zugänglichen, stoffeigenen, dimensionslosen *kritischen Parameters* $\alpha_k = \mathrm{d}\ln p/\mathrm{d}\ln T$ kann aber das thermische Verhalten nahezu aller Gase, zwischen deren Teilchen nur Dispersions- und Dipolkräfte wirken, gut beschrieben werden. Dieses *erweiterte Korrespondenzprinzip* ist deshalb zur Bestimmung unbekannter thermischer Daten reiner Gase für praktische Zwecke sehr nützlich.

Die Abweichungen im Verhalten realer Gase gegenüber dem idealen Zustand können auch mit Hilfe des *Kompressibilitäts-* oder *Realfaktors* $z = pV/RT$ beschrieben werden. Dieser dimensionslose Faktor ist für ideale Gase stets gleich 1, für reale Gase in der Regel verschieden von 1. Nach dem erweiterten Korrespondenzprinzip sollte z eine universelle Funktion von π, τ und α_k sein. Tatsächlich ergeben sich beim Auftragen von z über π für verschiedene Werte von τ Kurvenscharen, mit denen das Verhalten sehr vieler Gase einheitlich dargestellt und mit guter Näherung vorherbestimmt werden kann.

6.2.3. Zustandsänderungen bei der Kompression und Expansion

6.2.3.1. Abhängigkeit der inneren Energie vom Volumen. Beim Ausdehnen eines *idealen Gases* in ein Vakuum tritt keine Temperaturänderung auf und das Gas leistet auch keine Arbeit. Nach dem 1. Hauptsatz der Thermodynamik (↑ 6.1.2.2.) sind die innere Energie und die Enthalpie idealer Gase damit unabhängig vom Volumen und vom Druck, $(\partial u/\partial v)_T = 0$, $(\partial h/\partial p)_T = 0$. Diese auch ohne direkte experimentelle Begründung aus dem 2. Hauptsatz mit Hilfe der Gleichung des idealen Gases ableitbare Aussage bezeichnet man als *2. Gay-Lussacsches Gesetz*.

Bei der Entspannung *realer Gase* in ein Vakuum, d. h. ohne Leistung von Arbeit, tritt dagegen immer eine Temperaturänderung auf, die durch eine Energiezufuhr oder -abfuhr kompensiert werden muß, wenn der Vorgang isotherm ablaufen soll. Der Volumenkoeffizient der inneren Energie und der Druckkoeffizient der Enthalpie sind also bei realen Gasen nicht Null.

Dehnt sich ein Gas gegen eine äußere Kraft aus, indem es z. B. den Kolben eines Zylinders gegen den äußeren Luftdruck p verschiebt, so leistet es eine Volumenarbeit $\mathrm{d}w_{\mathrm{Vol}} = -p\,\mathrm{d}v$. Umgekehrt wird bei einer Kompression die Arbeit $\mathrm{d}w_{\mathrm{Vol}}$ gegen den Gasdruck p geleistet. Für *isobare* Volumenänderungen gilt also $w_{\mathrm{Vol}} = -p(v_2 - v_1)$. Erfolgt die Ausdehnung oder Kompression *isotherm*, dann bleibt bei *idealen Gasen* die innere Energie konstant, es erfolgt nach dem 1. Hauptsatz eine quantitative Umwandlung von Wärme in Volumenarbeit bzw. umgekehrt. Mit $p = nRT/v$

ergibt sich für die *Arbeit* $w_{Vol} = -nRT \ln(v_2/v_1) = -nRT \ln(p_1/p_2)$. Sie ist danach ausschließlich durch das Verhältnis des Endvolumens v_2 zum Ausgangsvolumen v_1 bzw. des Ausgangsdrucks p_1 zum Enddruck p_2 bestimmt, nicht aber von den absoluten Volumina bzw. Drücken abhängig. Je größer dieses *Kompressionsverhältnis* ist, um so größer ist auch die zu gewinnende Arbeit. Dieser Tatbestand wird beispielsweise im Preßlufthammer genutzt. Mit *realen Gasen* ist die bei einer isothermen Expansion durch gleiche Wärmezufuhr zu gewinnende Arbeit immer kleiner als bei idealen Gasen.

6.2.3.2. Adiabatische Zustandsänderungen. Theoretisch wie praktisch gleich bedeutsam sind Zustandsänderungen von Gasen, bei denen kein Wärmeaustausch mit der Umgebung erfolgt.

Mit $dq = 0$ gilt für solche *adiabatischen Vorgänge* $dw_{Vol} = du$, jede Arbeit vom oder am Gas erfolgt auf Kosten oder zugunsten seiner inneren Energie.

6.2.3.2.1. *Kompression und Expansion idealer Gase.* Erhöhung der inneren Energie durch adiabatische Kompression, wie z. B. beim Aufpumpen eines Fahrradschlauchs, führt zu einer Erwärmung des Gases, adiabatische Expansion zu einer Abkühlung. Der quantitative Zusammenhang zwischen Arbeit und Temperaturänderung lautet bei einem idealen Gas $w_{Vol} = c_v(T_2 - T_1)$. Bei adiabatischen Zustandsänderungen idealer Gase tritt an Stelle des Boyle-Mariotteschen Gesetzes die *Poissonsche Gleichung* $pv^{\varkappa} = \text{const.}$ $\varkappa = c_p/c_v$ ist das Verhältnis der Wärmekapazitäten bei konstantem Druck und bei konstantem Volumen und heißt *Poissonscher Koeffizient.* Da er stets größer als 1 ist (↑ 6.1.2.2.), verlaufen die Druck-Volumen-Kurven, die *Adiabaten,* immer steiler als die

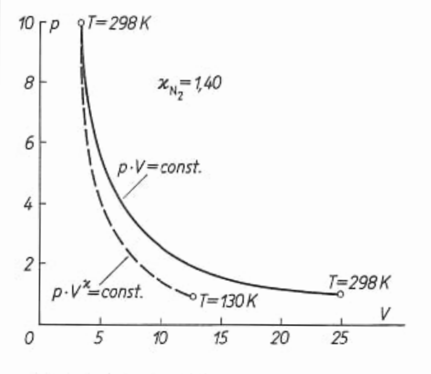


Abb. 6.2.3-1 Isotherme und Adiabate der Expansion von N_2 von 10 auf 10^5 Pa bei einer Ausgangstemperatur von 298 K; Druck p in 10^5 Pa, Molvolumen V in $1\ \text{mol}^{-1}$

Isothermen (Abb. 6.2.3-1). Die Poissongleichung ist für die Theorie vieler natürlicher und technischer Prozesse bedeutsam, z. B. für die Arbeitsvorgänge in Verbrennungsmotoren, für die Schallfortpflanzung in Gasen oder für die Bewegung großer Luftmassen in der Atmosphäre.

6.2.3.2.2. *Adiabatische Expansion realer Gase.* Sie verursacht infolge der Abhängigkeit der inneren Energie vom Volumen eine Temperaturänderung. Diese Erscheinung wurde von JOULE und THOMSON entdeckt. Mit Hilfe dieses sogenannten *Joule-Thomson-Effekts* ist es prinzipiell möglich, reale Gase bis zu ihrem Kondensationspunkt abzukühlen und damit zu verflüssigen.

Läßt man das Gas durch eine Drossel von einem höheren zu einem niedrigeren Druck expandieren, bleibt die Enthalpie dabei konstant, was als *isenthalpischer Drosseleffekt* bezeichnet wird und es tritt eine Temperaturänderung $\delta = (\partial T/\partial p)_h = -(1/c_p)\,(\partial h/\partial p)_T$ auf. Der *differentielle Joule-Thomson-Koeffizient* δ ist bei idealen Gasen stets Null; bei realen Gasen nimmt er bei einer von der Gasart abhängigen Temperatur T_I den Wert Null an, oberhalb dieser Inversionstemperatur ist er negativ, unterhalb positiv.

Unter der Bedingung $T > T_I$ erwärmt sich also das Gas bei einer adiabatischen Expansion, bei $T < T_I$ kühlt es sich dagegen ab. Eine Erklärung dieses Verhaltens liefert die Molekulartheorie der Gase. Danach sind oberhalb T_I die Abstoßungskräfte zwischen den Gasteilchen vorherrschend, unterhalb die Anziehungskräfte, die bei der Expansion auf Kosten der kinetischen Energie der Teilchen überwunden werden müssen. Hieraus ist auch verständlich, daß T_I für jede Gasart unterschiedlich ist und δ stark vom Druck und von der Temperatur abhängt. Mit sinkender Temperatur nimmt δ stark zu, die Abkühlung ist also um so stärker, je niedriger bereits die Ausgangstemperatur ist; mit steigendem Druck nimmt δ ab. Die Inversionstemperatur für He ist etwa 40 K, für H_2 222 K, für alle anderen Gase liegt sie oberhalb Normaltemperatur.

Bei der großtechnischen Verflüssigung der Luft oder anderer Gase nach dem *Lindeverfahren* arbeitet man mit Druckunterschieden von 1 bis $20 \cdot 10^7$ Pa. Dabei wird außerdem bereits entspanntes und abgekühltes Gas im Gegenstromprinzip zur Vorkühlung weiteren hochkomprimierten Gases verwendet. Auf diese Weise gelangt man nach einigen Kompressions-Expansions-Zyklen zu solchen Temperaturen, bei denen nach einer weiteren Expansion ein Teil des Gases kondensiert. H_2 und He erfordern aber zunächst eine Vorkühlung unter ihre Inversionstemperaturen, ehe man sie derart verflüssigen kann.

6.2.4. Molekulartheorie der Gase

Die *kinetische* oder *statistische Gastheorie* beschreibt die makroskopischen Zustandseigenschaften der Gase auf der Grundlage der Eigenschaften und Verhaltensweisen der Atome und Moleküle und zieht umgekehrt aus den Eigenschaften der makroskopischen Teilchensysteme (↑ 6.1.1.) Rückschlüsse auf die Einzelteilchen. Dazu werden die Methoden und Ergebnisse der *Molekularstatistik* (↑ 2.4.4.) auf geeignete Modelle für gasförmige Vielteilchensysteme angewandt.

6.2.4.1. Molekülmodelle. Im *Modell des idealen Gases* werden die Teilchen als starre Kugeln von bestimmtem Durchmesser und bestimmter Masse angesehen, die sich völlig regellos in einem Raum bewegen, gegen den ihr Eigenvolumen vernachlässigbar klein ist. Bei Zusammenstößen untereinander oder Stößen auf die Gefäßwände verhalten sie sich ideal-elastisch, wie Billardkugeln; zwischen ihnen wirken keinerlei Kräfte außer denjenigen, die bei den Zusammenstößen auftreten.

Zur besseren Annäherung an das reale Verhalten sind eine Reihe komplizierterer Modelle entwickelt und angewandt worden. Der Wirklichkeit am nächsten kommt das *Modell symmetrischer Moleküle mit zentral gerichteten Wechselwirkungskräften.* In ihm wird die Vorstellung harter Kugeln durch das Bild begrenzt deformierbarer Teilchen ersetzt, die auf andere sowohl zentral gerichtete anziehende als auch abstoßende Kräfte ausüben können. Ein solches Modell ist vollständig durch den Verlauf der potentiellen Wechselwirkungsenergie $\varphi(r)$ zweier Moleküle als Funktion ihres gegenseitigen Abstands r charakterisiert. Häufig angewandt wird das *Lennard-Jones-12 : 6-Potential* $\varphi(r) = 4\varepsilon[(\sigma/r)^{12} - (\sigma/r)^6]$.

ε und σ sind teilcheneigene und weitgehend temperaturunabhängige Konstanten, die experimentell, z. B. aus den Virialkoeffizienten (↑ 6.2.2.2.2.), bestimmt werden können. Das Glied mit negativem Vorzeichen entspricht einer Anziehung, während die Abstoßung entsprechend der hohen Potenz des ersten Glieds erst bei starker gegenseitiger Annäherung der Teilchen wirksam wird.

6.2.4.2. Klassische kinetische Theorie des idealen Gases. Sowohl der Druck eines Gases als auch die Temperatur sind ein Ausdruck der Teilchenbewegung. Der Gasdruck wird hervorgerufen durch den Aufprall der umherfliegenden Atome oder Moleküle auf die Behälterwände. Die Temperatur ist ein Maß für die mittlere kinetische Energie der Translation der Teilchen.

Für das Gesamtsystem und alle seine Teilchen sollen die Gesetzmäßigkeiten der klassischen Mechanik, z. B. der Energie- und der Impulserhaltungssatz gelten. Die ständigen Veränderungen der Bewegungsgeschwindigkeiten und -richtungen der Einzelteilchen durch die Zusammenstöße machen es jedoch erforderlich, zur Ableitung von Zusammenhängen zwischen der Summe der individuellen Bewegungsgrößen aller Teilchen und den makroskopischen Zustandseigenschaften des Gases die Erkenntnisse der Gleichgewichtsstatistik anzuwenden. Aussagen über die Verteilung der Geschwindigkeiten und Energien auf die einzelnen Teilchen eines Systems machen der *Maxwellsche Geschwindigkeitsverteilungssatz* und der *Boltzmannsche Energieverteilungssatz* (↑ 2.4.4.). Von grundlegender Bedeutung ist weiterhin der *Satz von der Gleichverteilung der Energie*, das *Äquipartitionsprinzip.* Der Satz besagt, daß eine von einem Teilchensystem aufgenommene Energie zu gleichen Anteilen auf alle Freiheitsgrade der Bewegung der Atome oder Moleküle verteilt wird. Auf jede der drei möglichen Bewegungskomponenten der Translation in den drei Raumrichtungen entfällt z. B. der gleiche Energiebetrag von $(1/2)\,k_B T$ (k_B Boltzmann-konstante).

6.2.4.2.1. *Gasdruck, Teilchengeschwindigkeiten und Temperatur.* Die Anwendung des Impulserhaltungssatzes auf ein Teilchensystem mit N Teilchen der Einzelmasse m in einem würfelförmigen Behälter von 1 cm³ ergibt für den Druck p unter Berücksichtigung der Geschwindigkeitsverteilung $p = \frac{1}{3} Nm\overline{w^2}$. $\overline{w^2}$ ist der arithmetische Mittelwert der Quadrate aller Teilchengeschwindigkeiten und Nm die Gasdichte. Mit $N = N_A/V$ (N_A Avogadro-Konstante) und der molaren Masse $M = N_A m$ folgt $pV = \frac{1}{3} M\overline{w^2} = \frac{2}{3}\,\frac{1}{2}\,M\overline{w^2} = \frac{2}{3}\bar{E}_{tr}$.

Das Produkt pV ist also der mittleren molaren Translationsenergie $\bar{E}_{tr}$ direkt proportional. Nach dem Gleichverteilungssatz muß andererseits $E_{tr} = 3\frac{1}{2} kTN_A = \frac{3}{2} RT$ sein. Damit ergibt sich aus der kinetischen Gastheorie die Gleichung des idealen Gases.

Für den aus dem Maxwellschen Geschwindigkeitsverteilungssatz folgenden Mittelwert der Quadrate der Teilchengeschwindigkeiten $\overline{w^2}$ liefern die angeführten Zusammenhänge die Beziehung $(w^2)^{1/2} = (3\,RT/M)^{1/2} = 157{,}93\,(T/M)^{1/2}\,\mathrm{ms^{-1}}$. Die hiernach für 25 °C berechneten Werte, z. B. 1921 ms⁻¹ für H_2, 482 ms⁻¹ für O_2, 515 ms⁻¹ für N_2 und 411 ms⁻¹ für CO_2 stimmen gut mit Ergebnissen experimenteller Untersuchungen an Molekülstrahlen überein. Aus der Theorie lassen sich noch zwei weitere Angaben fü. die Teilchengeschwindigkeiten ableiten. Die mittlere Geschwindigkeit $\bar{w}$ ergibt sich zu $\bar{w} = (8RT/\pi M)^{1/2} = 145{,}51\,(T/M)^{1/2}\,\mathrm{ms^{-1}}$; sie ist also um etwa

7,9 % kleiner als $\overline{w^2}$. Die wahrscheinlichste, d. h. am häufigsten vorkommende Geschwindigkeit $w' = (2RT/M)^{1/2} = 128{,}95\,(T/M)^{1/2}\,\mathrm{ms^{-1}}$ ist noch um etwa 11,4 % kleiner als $\bar{w}$.

Nach $T = \frac{2}{3}(\bar{E}_{tr}/R)$ ist die Temperatur der mittleren molaren Translationsenergie direkt proportional. Da $\bar{E}_{tr}$ nicht negativ werden kann, sind Temperaturen unterhalb 0 K unmöglich. Bei gleicher Temperatur müssen die Teilchen verschiedener Gase die gleiche mittlere molare Translationsenergie besitzen; die mittleren Geschwindigkeiten dagegen sind, wie die angegebenen Zahlenwerte zeigen, unterschiedlich groß.

6.2.4.2.2. *Stoßzahl und mittlere freie Weglänge.* Trotz der hohen Translationsgeschwindigkeiten bewegen sich die Teilchen eines Gases nur relativ langsam in eine *bestimmte* Richtung, da sie durch die gegenseitigen Zusammenstöße ständig ihre Bewegungsrichtung ändern. Die *Stoßzahl* Z, d. h. die Zahl der Zusammenstöße, die ein kugelförmiges Teilchen vom Druchmesser d in einer Sekunde erfährt, hängt vom *Stoßquerschnitt* πd^2, von der Teilchenkonzentration N (Zahl der Teilchen in 1 $\mathrm{cm^3}$) und der mittleren Teilchengeschwindigkeit $\bar{w}$ ab. Es gilt daher unter Berücksichtigung der Relativgeschwindigkeiten der Teilchen zueinander und der Geschwindigkeitsverteilung $Z = \sqrt{2}\pi d^2 \bar{w} N$. Unter Normalbedingungen hat Z bei vielen Gasen Werte in der Größenordnung von $5 \cdot 10^9$ bis $10^{10}\,\mathrm{s^{-1}}$ (Tab. 6.2.4-1). Die Gesamtzahl aller Zusammenstöße, die in 1 $\mathrm{cm^3}$ eines Gases in 1 s stattfinden, ist $ZN/2$; jedes der N Teilchen erfährt Z Stöße, an jedem Stoß sind aber zwei Teilchen beteiligt.

Die Strecke, die ein Teilchen im Mittel zwischen zwei aufeinanderfolgenden Stößen zurücklegt, heißt *mittlere freie Weglänge* $\bar{l} = w/Z = 1/(\sqrt{2}\pi d^2 N)$; $\bar{l}$ ist dem Gasdruck umgekehrt proportional. Für die meisten Gase ergeben sich damit unter Normalbedingungen mittlere freie Weglängen um 10^{-5} cm (↑ Tab. 6.2.4-1). Beide Größen sind von grundsätzlicher Bedeutung für alle Transportvorgänge in Gasen (↑ 6.2.5.).

Tab. 6.2.4-1 *Stoßzahlen* Z in $\mathrm{s^{-1}}$ und *mittlere freie Weglängen* $\bar{l}$ in cm in verschiedenen Gasen bei Normalbedingungen ($p_0 = 1{,}01325 \cdot 10^5$ Pa, $T_0 = 273{,}15$ K)

Gasart	$Z \cdot 10^{-9}$	$\bar{l} \cdot 10^5$
N_2	7,58	0,600
O_2	6,57	0,647
H_2	15,14	1,123
CO_2	9,13	0,397
CO	7,78	0,584
Cl_2	9,95	0,287
NH_3	13,23	0,441
He	6,69	1,798
Ne	4,32	1,238
Ar	5,99	0,635

6.2.4.2.3. *Effusion und Transfusion.* Die unterschiedlichen Geschwindigkeiten verschieden schwerer Atome oder Moleküle lassen sich direkt beim Ausströmen der Gase in ein Vakuum durch eine Düse nachweisen, deren Durchmesser klein im Vergleich zur mittleren freien Weglänge ist. Bei diesem als *Effusion* bezeichneten Vorgang wird die Geschwindigkeitsverteilung der im Gefäß vorhandenen Teilchen nicht beeinflußt. Nach dem *Graham-Bunsenschen Gesetz* verhalten sich daher die Effusionsgeschwindigkeiten w_{E1} und w_{E2} zweier Gase bei konstanter Temperatur wie die mittleren Geschwindigkeiten $\bar{w}_1$ und $\bar{w}_2$ ihrer Teilchen und sind damit umgekehrt proportional den Wurzeln aus den molaren Massen M_1 bzw. M_2: $w_{E1}/w_{E2} = \bar{w}_1/\bar{w}_2 = (M_2/M_1)^{1/2}$.

Die Tatsache, daß leichtere Gase schneller ausströmen als schwerere, ermöglicht die Trennung von Gasgemischen. An Stelle einzelner Düsen werden dazu poröse Scheidewände benutzt, deren mittlerer Porendurchmesser wiederum klein gegen $\bar{l}$ sein muß. Die Theorie der *Transfusion* mit Hilfe geeigneter Trennstufen wurde 1932 von HERTZ ausgearbeitet. Am bekanntesten ist die Anwendung eines entsprechenden Verfahrens zur großtechnischen Trennung der Isotope U-235 und U-238 in Form des gasförmigen UF_6 geworden. Als *Knudsensche Effusionsmethode* hat diese Trennmöglichkeit in neuerer Zeit steigende Bedeutung für die analytische Untersuchung von Gasgemischen, insbesondere bei höheren Temperaturen, erlangt. Die entsprechend ihrer molaren Masse verschieden schnell aus einer Düse in ein Hochvakuum ausströmenden Bestandteile werden in einem angeschlossenen Massenspektrometer identifiziert.

6.2.4.3. Statistische Theorie der Molwärmen. Aus der Molekulartheorie lassen sich auch Aussagen über die innere Energie der Gase und ihre Temperaturabhängigkeit, d. h. nach $(\partial U/\partial T)_v = C_v$ über ihre Wärmekapazitäten bzw. Molwärmen (↑ 6.1.2.2.2.) gewinnen. Grundlage dafür ist wiederum der Gleichverteilungssatz der Energie. Auf jeden Freiheitsgrad der Translation und der Molekülrotation entfällt die Energie $1/2 k_B T$, auf jeden Freiheitsgrad der Molekülschwingungen $2 \cdot 1/2 k_B T$, da bei jeder Schwingung ein Anteil kinetischer und ein Anteil potentieller Energie unabhängig voneinander zu berücksichtigen ist.

Die innere thermische Energie eines idealen Gases setzt sich aus der Summe der Energien aller Bewegungsformen zusammen, $U = \bar{E}_{tr} + \bar{E}_{rot} + \bar{E}_{schw}$. Bei sehr hohen Temperaturen

kann noch ein weiterer Energieanteil von der Elektronenanregung hinzukommen. Für eine Stoffmenge von 1 Mol eines idealen Gases resultiert mit f_{rot} = Zahl der Rotationsfreiheitsgrade und f_{schw} = Zahl der Schwingungsfreiheitsgrade $U = 3RT/2 + f_{rot}RT/2 + f_{schw}RT$. Für die Molwärme müßte danach gelten $C_v = 3/2R + 1/2f_{rot}R + f_{schw}R$. Die Molwärme C_v *einatomiger idealer Gase* sollte wegen $f_{rot} = f_{schw} = 0$ stets den gleichen, temperaturunabhängigen Wert $3R/2 = 12{,}47\ JK^{-1}mol^{-1}$ haben. Dieses Ergebnis konnte bei vielen Metalldämpfen und den Edelgasen tatsächlich auch experimentell bestätigt werden.

Zweiatomige Moleküle besitzen 3 Translations-, 2 Rotations- und einen Schwingungsfreiheitsgrad, die Valenzschwingung der beiden Atome gegeneinander. Die Molwärme sollte daher $7R/2 = 29{,}10\ JK^{-1}mol^{-1}$ sein. Im Experiment nähern sich die C_v bei hohen Temperaturen tatsächlich diesem Wert, sie sind aber temperaturabhängig und konvergieren bei niedrigen Temperaturen gegen $5R/2$. Eine Ausnahme ist H_2 mit einem Wert von konstant $5R/2$, der unterhalb Normaltemperatur auf $3R/2$ abfällt.

Sieht man bei *mehratomigen Molekülen* zunächst von den Schwingungen ab, so sollten derartige Teilchen bei linearem (gestrecktem) Aufbau 3 Translations- und 2 Rotationsfreiheitsgrade und eine Molwärme von $C_v = 5R/2 = 20{,}79\ JK^{-1}mol^{-1}$ oder bei gewinkeltem Aufbau (3 Rotationsfreiheitsgrade) von $C_v = 6R/2 = 24{,}94\ JK^{-1}mol^{-1}$ aufweisen. Tatsächlich konvergieren die experimentellen Ergebnisse bei niedrigen Temperaturen auf diese Werte. CO_2 hat demnach einen gestreckten Aufbau. Auffällig ist bei beiden Gruppen die starke Temperaturabhängigkeit, die im allgemeinen um so größer ist, je mehr Atome das Molekül aufbauen.

Moleküle mit n Atomen haben, wenn sie gestreckt sind, $3n - 5$, wenn sie gewinkelt aufgebaut sind, $3n - 6$ Schwingungsfreiheitsgrade. Alle zusammengesetzten innermolekularen Schwingungsbewegungen lassen sich auf diese *Normalschwingungen* zurückführen. Die experimentellen Werte führen zu dem Schluß, daß bei tiefen Temperaturen Schwingungen offenbar keinen Beitrag zur Teilchenenergie liefern und erst mit steigender Temperatur nach und nach angeregt werden. Daß Molekülschwingungen bei tiefen Temperaturen „eingefroren" sein sollen, steht allerdings im Widerspruch zum Gleichverteilungssatz und damit zu den Annahmen der klassischen Molekularstatistik.

Immerhin vermag diese Theorie aber die experimentellen Werte für ein- und zweiatomige Teilchen zumindest bei niedrigen bzw. sehr hohen Temperaturen gut zu erklären. Mit ihrer Hilfe ist es auch möglich, die Grenzwerte der Molwärmen C_v unter Berücksichtigung der Molekülschwingungen für Gase im idealen Zustand bei sehr hohen Temperaturen zu berechnen. Meist werden diese Grenzwerte bei der experimentellen Untersuchung jedoch nicht erreicht, da die Moleküle bereits thermisch dissoziieren, ehe alle Schwingungen voll angeregt sind.

6.2.4.4. Quantentheorie der Molwärme. Die Deutung der Temperaturabhängigkeit der Molwärmen bei niedrigen und mittleren Temperaturen durch allmähliche Anregung der Molekülschwingungen ist auf der Grundlage der *Quantentheorie* (↑ 2.5.) möglich. Danach können schwingungsfähige molekulare Teilchen Energie nicht in beliebigen Beträgen aufnehmen oder abgeben, sondern nur in Form ganzzahliger Vielfacher der Energiequanten $\varepsilon = h\nu$ ($h = 6{,}626 \cdot 10^{-34}$ Js Plancksches Wirkungsquantum; ν Frequenz der betreffenden Schwingung).

Bei tiefen Temperaturen ist die nach der klassischen Theorie errechnete mittlere Schwingungsenergie k_BT je Schwingungsfreiheitsgrad eines Moleküls klein gegenüber $h\nu$; dann wird praktisch überhaupt keine Anregung erfolgen. Mit Steigerung der Temperatur können zunächst einzelne, dann auch mehrere Energiequanten aufgenommen werden, d. h. die Schwingungen werden allmählich immer stärker angeregt. Die Häufigkeit von Molekülen mit den Energiewerten $\varepsilon = h\nu$, $2h\nu$, $3h\nu$ usw. in Abhängigkeit von der Temperatur ist durch den Boltzmannschen Energieverteilungssatz gegeben. Durch die quantenhafte Anregung der Schwingungen wird daher bei mittleren Temperaturen der tatsächliche Beitrag der Schwingungsenergien zur Molwärme stets kleiner sein als nach der klassischen Theorie. Erst bei hinreichend hohen Temperaturen wird $k_BT \gg h\nu$; dann sind praktisch alle Schwingungen angeregt, die Tatsache der Quantelung der Energie spielt keine Rolle mehr, es gilt der klassische Gleichverteilungssatz und beide Theorien ergeben die gleichen Grenzwerte.

Die Energiequanten der Teilchentranslation und im allgemeinen auch der Molekülrotationen sind so klein, daß diese Bewegungsformen schon bei tiefen Temperaturen praktisch immer vollständig angeregt sind. Eine Rotation um die Verbindungslinie der Atomkerne braucht dabei wegen des großen Trägheitsmoments bei zwei- und mehratomigen linearen Molekülen in der Regel nicht berücksichtigt zu werden. Nur beim H_2-Molekül mit seinem extrem kleinen Trägheitsmoment auch für Rotationen um senkrecht auf der Kernverbindungslinie stehende Achsen, d. h. relativ großen Quanten der Rotation, friert bei niedrigen Temperaturen die Rotationsbewegung insgesamt ein. Erst bei Temperatursteigerung auf etwa 300 K übersteigt die thermische Energie die Größe der zugehörigen

Rotationsquanten, was sich im Anstieg von C_v von $3R/2$ auf $5R/2$ ausdrückt.
Alle angedeuteten Sachverhalte lassen sich mit Hilfe der Quantenstatistik quantitativ formulieren und auswerten, so daß auf diesem Wege eine Berechnung der Molwärmen idealer Gase aus ihren Translations-, Rotations- und Schwingungsanteilen möglich ist. Alle dazu erforderlichen Moleküldaten, insbesondere Schwingungsfrequenzen, lassen sich, zumindest bei einfachen Molekülen, aus den Molekülspektren (↑ 5.3.) gewinnen.

6.2.5. Transporterscheinungen in Gasen

Die Erscheinungen der *Viskosität*, der *Wärmeleitung* und der *Diffusion* in Gasen beruhen auf der im Einzelnen völlig regellosen thermischen Molekularbewegung. Sie beinhalten einen Transport von Impulsen, Energien oder Massen in Richtung eines Geschwindigkeits-, Temperatur- oder Konzentrationsgefälles und sind deshalb als *statistische Ausgleichsvorgänge* aufzufassen. Auf einem Energietransport anderer Art beruht auch die *Schallausbreitung*.

6.2.5.1. Innere Reibung. Bei der Bewegung einer Gasschicht längs einer anderen sind beide durch die Teilchen, die infolge ihrer thermischen Bewegung von einer Schicht in die benachbarte gelangen, gewissermaßen miteinander verzahnt. Infolge des wechselseitigen Übertritts, der einer Übertragung verschieden großer Impulse I gleichkommt, versuchen sich die unterschiedlichen Bewegungsgeschwindigkeiten der Schichten auszugleichen. Zur Aufrechterhaltung eines Geschwindigkeitsunterschieds ist daher die Aufwendung einer die Reibungskraft $F_R = dI/dt$ kompensierenden Kraft erforderlich. Der in einer Sekunde übertragene Impuls ist der Reibungsfläche A und dem Geschwindigkeitsgefälle dw/dx senkrecht zur Bewegungsrichtung proportional, $F_R = \eta A(dw/dx)$. Der Proportionalitätsfaktor η in diesem *Newtonschen Fließgesetz* für laminare Strömung ist eine nur von der Temperatur abhängige Stoffkonstante. Er kann experimentell bestimmt werden und wird als *dynamischer Viskositätskoeffizient* oder kurz *dynamische Viskosität* bezeichnet. η ist ein Maß für die Größe der Schubspannung F_R/A ($\text{mkgs}^{-2}\text{m}^{-2}$), die zum Hervorrufen eines Geschwindigkeitsgefälles $dw/dx = 1$ ($\text{ms}^{-1}\text{m}^{-1}$) aufgewandt werden muß. Unter Normalbedingungen liegen die Viskositäten der Gase in der Größenordnung von $10^{-5}\ \text{kgm}^{-1}\text{s}^{-1} = 10^{-5}\ \text{Pa}\cdot\text{s}$ (Tab. 6.2.5-1). Die klassische kinetische Gastheorie ergibt unter Berücksichtigung der Geschwindigkeitsverteilung für die *Viskosität idealer Gase* $\eta = 1/2\,(M/V)\,\overline{w}\overline{l}$. Sie ist damit über $\overline{w}$ der Wurzel aus der Temperatur proportional (↑ 6.2.4.2.1.), bemerkenswerterweise jedoch unabhängig vom Druck, was von MAXWELL bereits 1876 vorausgesagt wurde. Eine Änderung der Dichte M/V mit dem Druck wird gerade kompensiert durch die dabei auftretende gegensinnige Änderung der mittleren freien Weglänge $\overline{l}$. Diese theoretischen Ergebnisse konnten experimentell bestätigt werden.

Tab. 6.2.5-1 Dynamische Viskositätskoeffizienten η in $\text{kgm}^{-1}\text{s}^{-1}$ und Wärmeleitfähigkeiten λ in $\text{Wm}^{-1}\text{K}^{-1}$ verschiedener Gase bei Normalbedingungen ($p_0 = 1{,}01325 \cdot 10^5$ Pa, $T_0 = 273{,}15$ K)

Gasart	$\eta \cdot 10^5$	$\lambda \cdot 10^2$
Luft	1,71	2,44
N_2	1,67	2,43
O_2	1,93	2,47
H_2	0,85	17,42
CO_2	1,38	1,42
CO	1,66	2,34
CH_4	1,03	3,01
SO_2	1,18	8,37
NH_3	0,94	2,18
He	1,89	14,24
Ne	2,98	4,61
Ar	2,10	1,63

Die *Viskosität realer Gase* zeigt vor allem eine stärkere Temperaturabhängigkeit. Auf der Grundlage verbesserter Molekülmodelle wurden unter Berücksichtigung der Deformierbarkeit der Teilchen bei Zusammenstößen sowie der zwischenmolekularen Wechselwirkungskräfte halbempirische Beziehungen aufgestellt, von denen die *Sutherlandgleichung* mit den Stoffkonstanten B, der *Sutherlandkonstanten*, und C bei geeigneter Anpassung dieser Konstanten an Meßwerte die Viskositäten vieler Gase in einem größeren Temperaturbereich gut wiedergibt: $\eta = B\sqrt{T}/(1 + C/T)$.

Bei *höheren Gasdichten* wird die Viskosität auch vom Druck abhängig. Nach der zuerst von ENSKOG entwickelten Theorie muß dann neben dem *Impulstransport* durch die mehr oder weniger freie Bewegung der Teilchen zwischen den Zusammenstößen auch eine Impulsübertragung durch die Zusammenstöße selbst berücksichtigt werden.

6.2.5.2. Wärmeleitung. Der Energietransport in Richtung eines Temperaturgefälles erfolgt nach der Gleichung $dq/dt = -\lambda A(dT/dx)$. Die senkrecht durch eine Fläche A in einer Sekunde strömende Wärme dq/dt ist dieser Fläche und dem Temperaturgefälle dT/dx in der Strömungsrichtung proportional. Der Proportionalitätsfaktor λ ist eine temperaturabhängige Stoffkonstante und heißt *Wärmeleitfähigkeit* oder *Wärmeleitzahl*. λ wird in der Einheit $\text{Wm}^{-1}\text{K}^{-1}$

angegeben. Mit dem Minuszeichen in der Wärmeleitgleichung wird berücksichtigt, daß der Wärmestrom in Richtung der niedrigeren Temperatur fließt.

Unter Normalbedingungen liegen die Wärmeleitfähigkeiten der Gase in der Größenordnung von 10^{-2} bis 10^{-1} $Wm^{-1}K^{-1}$ (↑ Tab. 6.2.5-1). Diese gegenüber Flüssigkeiten und festen Stoffen relativ geringe thermische Leitfähigkeit wird vielfach praktisch genutzt, z. B. bei wollenen Kleidungsstücken, Bettdecken und Doppelfenstern. Sie ist auch die Erklärung für das sogenannte *Leidenfrostsche Phänomen*: Ein Wassertropfen verdampft auf einer glühenden Platte nicht sofort, sondern infolge der ihn umgebenden Dampfhaut erst nach einer gewissen Zeit.

Die klassische kinetische Theorie der idealen Gase liefert unter bestimmten Näherungsvoraussetzungen für einatomige Gase $\lambda = 1{,}26\,(M/V)\,c_{sp,v}\bar{w}\bar{l}$. M/V ist wiederum die Dichte, $c_{sp,v}$ die spezifische Wärme bei konstantem Volumen. Für zweiatomige Gase wird der Zahlenfaktor 0,95. Wie bei der Viskosität ist die Gültigkeit auf den idealen Zustand beschränkt.

Die inneren Zusammenhänge der Transporterscheinungen in verdünnten Gasen kommen u. a. darin zum Ausdruck, daß die Stoffkenngrößen η und λ ebenso wie der Diffusionskoeffizien D (↑ 6.2.5.3.) dem Produkt $\bar{w}\bar{l}$ proportional sind. Da die mittlere freie Weglänge $\bar{l}$ dem Druck umgekehrt proportional ist (↑ 6.2.4.2.2.), wird sie bei 10^{-1} Pa etwa 10^6mal größer als bei Atmosphärendruck und erreicht z. B. für Luft Werte in der Größenordnung von 10 cm. Befindet sich ein Gas von solch niedrigem Druck in einem Behälter von kleinen Abmessungen, dann bewegen sich die meisten Teilchen direkt von einer Gefäßwand zur anderen, ohne noch untereinander zusammenzustoßen. $\bar{l}$ wird damit im wesentlichen nicht mehr durch die Zahl der Zusammenstöße zwischen den Teilchen, sondern durch die Behälterabmessungen bestimmt und bleibt daher auch bei weiterer Senkung des Drucks praktisch konstant. In diesem sogenannten *Knudsen-Gebiet* sind dann aber η und λ nicht mehr vom Druck unabhängig, sondern nach den angeführten Beziehungen der Dichte, dem Druck proportional. Hierdurch wird verständlich, daß die Wärmeisolation in Thermosflaschen durch Evakuieren des Zwischenraums zwischen den beiden Glaswänden verbessert wird, wenn man ein Vakuum von mindestens 1 Pa oder besser erzeugt.

6.2.5.3. Diffusion. Als Diffusion bezeichnet man jeden von selbst vor sich gehenden *Massetransport*, der durch Konzentrations- (Partialdruck-) Differenzen zwischen verschiedenen Bereichen einer Mischphase verursacht wird. Er verläuft in geschlossenen Systemen stets bis zum völligen Konzentrationsausgleich, d. h. bis zur statistisch völlig gleichmäßigen Verteilung aller Komponenten in der Mischphase. Infolge der weitaus größeren Beweglichkeit aller Teilchen verläuft die Diffusion in Gasen etwa 10^5mal schneller als in Flüssigkeiten.

Grundgleichung der Diffusion ist das *1. Ficksche Diffusionsgesetz* $dn/dt = -DA\,(dc/dx)$. Die senkrecht durch eine Fläche A in der Zeiteinheit diffundierende Stoffmenge ist dieser Fläche und dem Konzentrationsgradienten $-dc/dx$ proportional. Das Minuszeichen bedeutet, daß die Diffusion stets in Richtung der niedrigeren Konzentration verläuft. Der Proportionalitätsfaktor D ist der *Diffusionskoeffizient*. Er wird gewöhnlich in der Einheit cm^2s^{-1} angegeben.

Unter Normalbedingungen liegen die Diffusionskoeffizienten von Gasen in der Größenordnung von 10^{-2} bis 1 cm^2s^{-1}. Sie hängen stark von der Natur der diffundierenden Teilchen, von der Art des Diffusionsmediums, vom Druck, d. h. von der Gasdichte, und von der Temperatur ab.

Die theoretische Behandlung der Diffusion ist selbst unter vereinfachenden Annahmen und mit den einfachsten Molekülmodellen kompliziert. Aus der klassischen kinetischen Theorie der idealen Gase ergibt sich z. B. für den sogenannten *Selbstdiffusionskoeffizienten*, den man in reinen Gasen durch radioaktiv markierte Atome oder Moleküle bestimmen kann, $D = 1/3\bar{w}\bar{l}$. D ist dem Druck umgekehrt proportional und bei konstantem Druck direkt proportional $T^{3/2}$. In Gasgemischen hängt D noch vom Mischungsverhältnis ab.

Erzeugt man in einer Gasmischung ein Temperaturgefälle, dann diffundieren die leichteren Teilchen etwas stärker in Richtung der höheren Temperatur. Das Ergebnis dieser *Thermodiffusion* ist eine Entmischung, die aber im stationären Zustand nur sehr geringfügig ist, da ihr das Bestreben zum Konzentrationsausgleich durch die normale Diffusion entgegenwirkt.

Das Gleichgewicht zwischen Thermodiffusion und Diffusion läßt sich in Richtung einer stärkeren Entmischung verschieben, wenn der Thermodiffusionseffekt durch eine Konvektionsströmung unterstützt wird. In dem von Clusius und Dickel 1938 entwickelten *Trennrohr*, einem langen, senkrecht aufgestellten Rohr mit Außenkühlung und einem axialen Heizdraht, reichern sich die leichteren Moleküle in der Rohrmitte etwas an. Sie bewegen sich durch die Wärmekonvektion längs des Drahts nach oben, während die schwereren Teilchen an der Rohrwandung nach unten gleiten. Durch dieses Zusammenspiel von Thermodiffusion und Konvektion im Gegenstromkreislauf reichern sich immer weitere Anteile des leichteren Gases

am heißen Draht und des schwereren an der Wandung an, der Trenneffekt vergrößert sich erheblich. Nach einer bestimmten Zeit können am oberen Ende des Drahts praktisch vollständig das leichtere Gas und am unteren Ende der Rohrwandung das reine schwerere Gas entnommen werden. Auf diese Weise lassen sich sogar Isotopengemische bis zu Reinheitsgraden über 99,9 % trennen.

6.2.5.4. Schallausbreitung. Die Ausbreitung des Schalls ist an das Vorhandensein eines stofflichen Mediums gebunden, dessen Teilchen einen *Energietransport* in Form von elastischen Longitudinalwellen, d. h. Kompressionswellen mit abwechselnder Kompression und Dilatation des Mediums, ermöglichen. Im Bereich des Hörschalls zwischen 16 Hz und etwa 20 kHz sowie im Ultraschallbereich bis etwa 100 MHz erfolgen die Druckänderungen in Gasen adiabatisch (↑ 6.2.3.2.).

Die *Schallgeschwindigkeit* in idealen Gasen ist gegeben durch $w_{Sch} = \sqrt{\varkappa p/\varrho}$. $\varkappa$ ist der Poissonkoeffizient c_p/c_v, p der Gasdruck und ϱ die Gasdichte. Da bei konstanter Temperatur die Dichte dem Druck proportional ist, hängt die Schallgeschwindigkeit nicht vom Druck ab. Mit der Temperatur nimmt sie dagegen proportional zu $\sqrt{T}$ zu, da nach der Gleichung des idealen Gases $pv = nRT$ und $\varrho = nM/v$ ist. Unter Normalbedingungen beträgt sie in trockener Luft 331,45 ms^{-1}.

Beim Lauf einer Schallwelle durch ein Gas nimmt die Schallintensität ständig ab, da die Schallenergie nach und nach von den Gasteilchen aufgenommen und irreversibel in innere Energie, z. B. in Form von Wärme, umgewandelt wird. Zur *Schallabsorption* tragen vor allem die Vorgänge der inneren Reibung, der Wärmeleitung und bei Gasgemischen auch der Diffusion bei. Dadurch hängt z. B. die Reichweite des Schalls in der Luft stark vom Wasserdampfgehalt ab.

6.3. Flüssigkeiten

6.3.1. Charakteristische Merkmale und Einteilung

Nach zunehmender Stärke der zwischenmolekularen Wechselwirkungen bzw. abnehmender Beweglichkeit der Teilchen geordnet, schließen sich an die Gase die Flüssigkeiten an.

Im *flüssigen Aggregatzustand* werden die Wirkungen der Teilchenbewegung und der Anziehung zwischen ihnen vergleichbar groß. Der gegenseitige Abstand zwischen den Atomen oder Molekülen schwankt daher nur wenig um einen Mittelwert, obwohl die Teilchen erhebliche *Translationsbewegungen* ausführen, allerdings stets im Kraftfeld zwischenmolekularer Wechselwirkungen (↑ 5.1.4.). Daraus ergeben sich die makroskopischen Eigenschaften der *Volumenbeständigkeit* bei relativ geringer Abhängigkeit von den Zustandsbedingungen sowie der unbegrenzten *Formveränderlichkeit* schon bei Einwirkung sehr geringer Kräfte. Abgesehen von Zustandsbereichen in der Nähe des kritischen Punkts (↑ 6.2.2.2.1.) haben die Flüssigkeiten in vielen Eigenschaften größere Gemeinsamkeiten mit festen Stoffen als mit Gasen. Sie bilden Grenzflächen aus, ihre Dichten sind relativ groß und die Kompressibilitäten und thermischen Ausdehnungen sind relativ gering; auch unterscheiden sich die Wärmekapazitäten ein und desselben Stoffs im festen und im flüssigen Zustand nur wenig voneinander. Flüssigkeiten und feste Stoffe faßt man daher auch unter der Bezeichnung *kondensierte Stoffe* zusammen.

Nach ihrem stofflichen Aufbau unterscheidet man zwei Gruppen von Flüssigkeiten. *Reine Flüssigkeiten* bestehen nur aus Teilchen gleicher stöchiometrischer Zusammensetzung und gegebenenfalls deren Assoziations- oder Dissoziationsprodukten. Nach der Art und Struktur ihrer Teilchen sind sie unpolare oder polare Flüssigkeiten. Die Atome oder Moleküle in *unpolaren Flüssigkeiten* werden im wesentlichen allein durch Dispersionskräfte zusammengehalten, z. B. in Br_2, CS_2, CCl_4, Alkanen. Zwischen den Teilchen der *polaren Flüssigkeiten* wirken dagegen auch beträchtliche elektrostatische Dipolwechselwirkungskräfte, z. B. in HCN, $CHCl_3$, CH_3OCH_3, oder gar Wasserstoffbrückenbindungen, wie in H_2O oder den Alkoholen. Zwei Sonderformen der reinen Flüssigkeiten sind die geschmolzenen Salze, sogenannte ionische Flüssigkeiten, und die flüssigen Metalle.

Flüssige Mischphasen sind homogene Gemische aus Teilchen mehrerer Stoffe. Sind diese in vergleichbarer Menge vorhanden, spricht man von *flüssigen Mischungen*. Überwiegt dagegen eine Teilchenart, das Lösungsmittel, gegenüber den anderen, den gelösten Stoffen, dann benutzt man die Bezeichnung *flüssige Lösungen*. Je nach der Anzahl der in ihnen enthaltenen Stoffe teilt man ein in *binäre*, *ternäre* usw. *Mischungen* oder *Lösungen*.

Infolge ihrer speziellen Eigenschaften werden zwei *Übergangsformen zum festen Aggregatzustand* gesondert betrachtet. *Kristalline Flüssigkeiten* sind ihrem makroskopischen Erscheinungsbild nach zwar flüssig, haben aber ausgeprägte strukturelle Merkmale des Kristallzustands (↑ 6.3.2.1.4.). Manche *unterkühlten Flüssigkeiten* sind dagegen feste Stoffe, z. B. die Gläser, denen aber das charakteristische Merkmal der Festkörper, die regelmäßige dreidimensionale Anordnung der Bausteine in einem Raumgitter, weitgehend fehlt (↑ 6.5.3.).

6.3.2. Zustandseigenschaften reiner Flüssigkeiten

6.3.2.1. Struktur reiner Flüssigkeiten. Die Eigenschaften der Flüssigkeiten werden sehr wesentlich von den teilcheneigenen Wechselwirkungskräften bestimmt; infolge der thermischen Molekularbewegung kommt es jedoch nicht zur Ausbildung geordneter Strukturen über größere Bereiche, wie sie beim Festkörper vorliegen (↑ 6.4.1.). In Mikrobereichen existieren dagegen gewisse *Ordnungszustände.*
Im Unterschied zur *Fernordnung* in Festkörpern charakterisiert man die in Flüssigkeiten vorliegende Struktur mit dem Begriff *Nahordnung*, d. h., nur die nächsten Nachbarn eines beliebigen Teilchens zeigen hinsichtlich ihres Abstands und ihrer Orientierung gegenüber diesem Teilchen zu jedem Zeitpunkt eine deutliche Ordnung. Durch die Temperaturbewegung ist sie bei den übernächsten Nachbarn schon wesentlich weniger ausgeprägt und bereits in Entfernungen von wenigen Teilchendurchmessern überhaupt nicht mehr vorhanden.
Ein wesentliches Merkmal der Nahordnung ist ihr dynamischer Charakter. Infolge der starken Molekularbewegung liegt der Mittelpunkt eines bestimmten Teilchens weder im Raum noch gegenüber den Nachbarteilchen fest. Die geordneten Mikrobereiche befinden sich vielmehr ständig im Auf- und Abbau und sind nicht etwa als stabile Mikrokristallite aufzufassen.
6.3.2.1.1. *Experimenteller Nachweis und theoretische Deutung.* Die wichtigste experimentelle Methode zur Untersuchung von Flüssigkeitsstrukturen ist die Beugung von Röntgenstrahlen, wie sie vor allem zur Erforschung von Gitterstrukturen kristalliner Festkörper angewandt wird (↑ 6.4.2.). Weitere Methoden sind z. B. die Beugung von Neutronenstrahlen, die IR-, die Raman- oder die Neutronenspektroskopie (↑ 5.3.4.). In den Röntgenbeugungsdiagrammen von Flüssigkeiten ist zwar nicht eine solche Vielfalt scharfer Linien wie bei Kristallen erkennbar, jedoch treten deutlich einige Schwärzungsmaxima niedriger Ordnung auf (Abb. 6.3.2-1). Sie erscheinen um so ausgeprägter, je stärker die Wechselwirkungen in der Flüssigkeit sind und je näher sich diese an ihrem Erstarrungspunkt befindet. Mit steigender Temperatur werden die Beugungsdiagramme durch den steigenden Einfluß der Teilchenbewegung immer unschärfer.

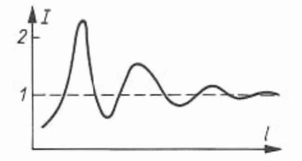


Abb. 6.3.2-1 Fotometrische Schwärzungskurve eines Röntgenbeugungsdiagramms von flüssigem Quecksilber (schematisch); I Schwärzungsintensität, l Entfernung vom Primärstrahl im Beugungsdiagramm; Ordinaten willkürlich

Die Höhen und Breiten der Schwärzungsmaxima entsprechen den Linienintensitäten und -schärfen im Beugungsdiagramm. Zu dessen Auswertung bestimmt man aus den Intensitätsverteilungen die *radialen Verteilungsfunktionen* $g(r)$. In einem Koordinatensystem mit dem Ursprung in einem betrachteten Teilchen stellt $g(r)$ die Wahrscheinlichkeit dar, den Schwingungsmittelpunkt eines benachbarten Teilchens am Ende eines vom Koordinatenursprung ausgehenden Vektors der Länge r anzutreffen. Die Maxima in Abb. 6.3.2-2b lassen für die ersten Nachbarn eines Teilchens tatsächlich eine Ordnung erkennen, die aber sehr schnell abklingt. Der Vergleich mit den radialen Verteilungsfunktionen für einen Festkörper (Abb. 6.3.2-2a) und für ein ideales Gas (Abb. 6.3.2-2c) läßt die typische Zwischenstellung der Flüssigkeiten besonders deutlich werden. Beim Kristall erreicht die Wahrscheinlichkeit für die Anwesenheit eines Teilchens periodisch über einen größeren Bereich immer wieder den gleichen Maximalwert, während sie beim Gas von einer bestimmten, durch den Teilchendurchmesser gegebenen Entfernung an überall gleich groß ist.
Die radiale Verteilungsfunktion liefert nur eine eindimensionale Information, nämlich über die Abstände zwischen den Teilchen, noch nicht aber über die Richtungen der Abstandsvektoren in der dreidimensionalen Struktur. Um dahin

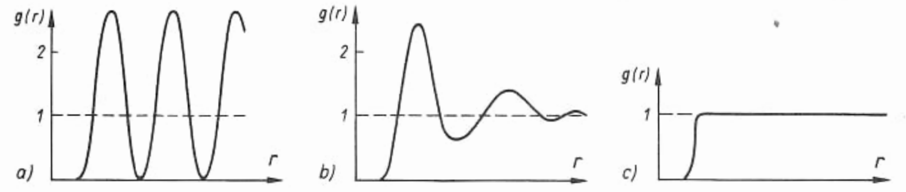


Abb. 6.3.2-2 Radiale Verteilungsfunktionen $g(r)$ als Darstellung der Ordnungszustände a) in idealen Kristallen, b) in Flüssigkeiten, c) in idealen Gasen

zu gelangen, müssen Modellvorstellungen zur Struktur entwickelt und die aus ihnen folgenden Aussagen auf Übereinstimmung mit den experimentellen Ergebnissen geprüft werden. Dies erfordert einen sehr hohen Rechenaufwand, der bis heute selbst mit den leistungsfähigsten elektronischen Rechenanlagen kaum zu bewältigen ist.

Zur molekulartheoretischen Deutung und Beschreibung der Nahordnung sind eine Vielzahl von *Strukturtheorien des flüssigen Zustands* aufgestellt worden. Eine geschlossene Theorie, die alle strukturabhängigen makroskopischen Stoffeigenschaften in sich konsistent zu beschreiben vermag, gibt es jedoch noch nicht. Dafür sind u. a. zwei Gründe maßgebend: Zum einen sind die Art und der strukturelle Aufbau der Bausteine sowie die Wechselwirkungen zwischen diesen außerordentlich mannigfaltig. Zum anderen weist der flüssige Zustandsbereich der Stoffe spezifische, nur ihm selbst eigene qualitative Merkmale auf, die sich aus einer Überlagerung wesentlicher Eigenschaften von Gasen und Festkörpern ergeben. Daher können z. B. alle *kinetisch-statistischen Strukturtheorien* auf der Basis der Theorie der dichten Gase ebenso wie alle *Theorien auf der Grundlage der Realstruktur von Kristallen* immer nur spezielle Flüssigkeitseigenschaften, gewissermaßen in einer einseitigen Beschreibungsweise einer komplexen Realität, annähernd genau wiedergeben bzw. erklären.

Erfolgversprechender erscheint dagegen der Weg, der von EYRING auf der Grundlage der früheren *Löchertheorie* gewiesen wurde und mit dem *gleichzeitig* die gasanalogen und die festkörperähnlichen Merkmale berücksichtigt werden. Die Kompressibilitätskoeffizienten der Flüssigkeiten sind wesentlich kleiner als die der Gase, aber immer noch etwa 100mal größer als die fester Stoffe. Unter Normalbedingungen hat nämlich jede Flüssigkeit etwa 3% leeres Volumen. Abb. 6.3.2-3 zeigt schematisch das Eyringsche Modell einer Flüssigkeit. Ein Gas besteht im wesentlichen aus leerem Volumen, in dem sich einige Teilchen regellos bewegen. Die flüssige Phase wird aufgefaßt als ein mit Teilchen gefülltes Volumen, in dem sich einige *Leerstellen, Löcher*, regellos bewegen. Die Größe der Löcher ist unterschiedlich und hängt von verschiedenen stoffeigenen und anderen Faktoren ab; unter bestimmten Bedingungen ergibt sich als Mittelwert etwa das Eigenvolumen eines Teilchens. Ein zu einem bestimmten Zeitpunkt unmittelbar an eine Leerstelle angrenzendes Teilchen verhält sich in gewisser Hinsicht wie ein Gasmolekül. Die Teilchen, die zu diesem Zeitpunkt dagegen gerade vollständig von anderen umgeben sind, haben demgegenüber die Eigenschaften von Bausteinen eines stark gestörten Kristalls. Aus diesen Vorstellungen können bestimmte Ansätze entwickelt werden, die manche Flüssigkeitseigenschaften bereits gut wiedergeben.

6.3.2.1.2. *Struktur des Wassers.* Besonders ausgeprägte Ordnungszustände sind in solchen Flüssigkeiten zu erwarten, deren Moleküle aufeinander relativ starke und gerichtete Wechselwirkungskräfte ausüben. Das beste Beispiel dafür ist der wichtigste und verbreitetste flüssige Stoff, das *Wasser*. Seine Allgegenwärtigkeit und die Selbstverständlichkeit des Umgangs mit ihm lassen kaum bewußt werden, daß es im Vergleich zu anderen Flüssigkeiten in fast allen physikalischen Eigenschaften große Anomalien aufweist. Dies gilt für das bekannte Dichtemaximum bei −3,98 °C ebenso, wie für die gegenüber anderen Flüssigkeiten ähnlichen Molekülaufbaus extremen Werte der Verdampfungswärme, der Oberflächenspannung, der Wärmeleitfähigkeit oder der Wärmekapazität sowie für die anomalen Temperaturabhängigkeiten der Kompressibilität und der Schallgeschwindigkeit. Alle diese Besonderheiten können nur durch spezielle Struktureigenschaften begründet sein.

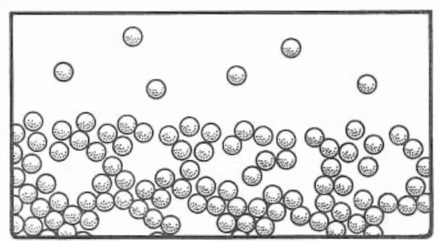


Abb. 6.3.2-3 „Momentaufnahme" einer Flüssigkeit mit ihrem Dampf nach dem Eyringschen Strukturmodell der Flüssigkeiten (schematisch)

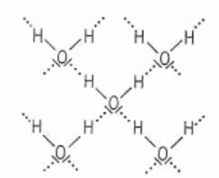


Abb. 6.3.2-4 Schematische Darstellung tetraedrischer Wasserstoffbrücken im Wasser

Eine Ausnahmestellung nimmt bereits das *Wassermolekül* H_2O ein. Es ist das einzige Molekül, das von einem einatomigen Zentrum aus nach allen vier Ecken eines Tetraeders über die beiden H-Atome und die beiden einsamen Elektronenpaare Wasserstoffbrücken ausbilden kann (Abb. 6.3.2-4). In Übereinstimmung mit Röntgenstrukturanalysen von Eis hat Wasser daher im festen Zustand einen hochsymmetrischen Aufbau. Im hexagonalen β-tridymitartigen Eisgitter sind die H_2O-Moleküle so angeordnet, daß jedes O-Atom tetraedrisch von

vier H-Atomen umgeben ist (Tafel 71). Dies führt zu einer sperrigen Gitterstruktur mit großen Hohlräumen, die sich in Richtung einer der kristallographischen Achsen zu Kanälen vereinigen und die relativ geringe Dichte des Eises erklären. Bei der Schmelztemperatur bricht das Gitter zusammen; das mit dem Eis im Gleichgewicht stehende Wasser hat eine größere Dichte, es müssen also – zumindest teilweise – die Hohlräume ausgefüllt sein.

Ausgangspunkt aller *Strukturtheorien des flüssigen Wassers* ist die auch experimentell durch Röntgenbeugungsaufnahmen und Ermittlung der radialen Verteilungsfunktionen nachgewiesene Tatsache, daß beim Zusammenbrechen des Eisgitters ein erheblicher Anteil der *Wasserstoffbrücken* erhalten bleibt. Auch im flüssigen Zustand werden deshalb kleinere Nahordnungsbereiche in Form tetraedrisch strukturierter, fluktuierender, *räumlicher Netzwerke*, sogenannter *Cluster*, angenommen, von denen ständig Einzelmoleküle abgespalten und andere an anderer Stelle wieder angelagert werden. Weitere Einzelmoleküle sind auf Zwischengitterplätzen in den Hohlräumen der β-Tridymitstruktur der dreidimensionalen Netzwerke eingelagert, deren Lebensdauer aber äußerst gering ist und die sich in ständigem Umbau befinden. Bei 0 °C liegen nur etwa 10% aller Moleküle als Einzelteilchen vor. Der Hauptteil ist in den Clustern mit größenordnungsmäßig etwa 100 Molekülen je Einheit enthalten. Diese werden mit steigender Temperatur unter gleichzeitiger weiterer Besetzung von Zwischengitterplätzen mit Monomeren, d. h. unter Erhöhung der Packungsdichte, abgebaut. Die Überlagerung dieses Vorgangs mit der thermischen Ausdehnung erklärt das Dichtemaximum und weitere Anomalien. Eine vollständige Deutung aller anomalen Eigenschaften des Wassers ist jedoch auch mit dem skizzierten Strukturmodell noch nicht möglich. Deshalb werden als dritte Strukturelemente kleinere, kompakte Molekülaggregate aus nur wenigen Teilchen als Übergangsformen zwischen den Netzwerken und den Einzelmolekülen vermutet. Der Abbau der Netzwerke bei Temperatursteigerung verläuft wahrscheinlich über diese Aggregate.

6.3.2.1.3. *Polywasser.* Experimentelle Untersuchungen der verschiedensten Art haben ergeben, daß Wasser in feinen Kapillaren bzw. in dünnsten Schichten unterhalb etwa 0,1 μm Dicke im Kontakt mit Festkörperoberflächen, aber z. B. auch mit biologischen Membranen, völlig abweichende Eigenschaftswerte gegenüber dem kompakten Stoff besitzt.

Solches Polywasser (Derjagin 1962) hat eine deutlich größere Dichte und Oberflächenspannung sowie eine bedeutend höhere Viskosität, und es erstarrt bei etwa −40 °C zu einer glasartigen Masse. Der Dampfdruck ist erheblich niedriger, und es beginnt daher erst bei etwa 150 °C zu sieden, wobei sich die anomalen Eigenschaften verstärken; erst oberhalb 700 °C zersetzt es sich vollständig zu normalem Wasser.

Die Möglichkeit, daß Verunreinigungen in den Kapillaren entweder von ihrer Herstellung her oder durch Auflösung von Bestandteilen der Silicatoberflächen die Ursachen für die beobachteten Eigenschaften des Kondensationsprodukts sein könnten, ist noch nicht auszuschließen, was bei der Bewertung neuerer Theorien zu bedenken ist.

6.3.2.1.4. *Kristalline Flüssigkeiten.* Bestimmte organische Verbindungen können innerhalb scharf begrenzter Temperaturbereiche in einem thermodynamisch stabilen Übergangszustand zwischen dem festen, kristallinen und dem flüssigen Aggregatzustand auftreten. Beim Schmelzen gehen diese Stoffe zunächst in einen *kristallin-flüssigen Zustand* über, in dem sie mehr oder weniger viskose und trübe Flüssigkeiten bilden. Erst bei einer höheren Temperatur wird die Schmelze klar durchsichtig. Dieser Umwandlungsvorgang ist reversibel.

Das Auftreten *thermotroper kristallin-flüssiger Zustandsbereiche* wird in erster Linie vom Molekülaufbau der betreffenden Stoffe bestimmt. Sie bestehen immer aus gestreckten, ebenen und weitgehend starren Molekülrümpfen. Eine häufige Bauform ist $R—C_6H_4—$ $—M—C_6H_4—R'$; die starre Mittelgruppe M, z. B. $—CH{=}CH—$, $—C{\equiv}C—$, $—N{=}N—$ oder $—CH{=}N—$, bildet mit den beiden Phenylengruppen ein konjugiertes aromatisches System und ist Sitz eines starken Dipolmoments. Die Flügelgruppen R und R' sind meist aliphatische Kohlenwasserstoff- und schwach

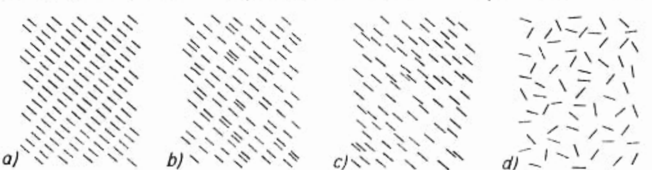


Abb. 6.3.2-5 Zweidimensionale Schemata der Ordnungszustände *a*) in kristallisierten Festkörpern mit speziellen langgestreckten Molekülen, b) in kristallinen Flüssigkeiten bei smektischen Phasen, c) bei nematischen Phasen, d) in isotropen Flüssigkeiten

polare Ether- oder Estergruppen; ihre Art und Kettenlänge bestimmen das Ordnungsprinzip der auftretenden kristallin-flüssigen Phasen und die Größe ihres Existenzbereichs.

Besonders aufgrund von permanenten und induzierten Dipolwechselwirkungen lagern sich die langgestreckten und stark polarisierbaren Moleküle auch noch oberhalb der Schmelztemperatur bevorzugt parallel zueinander. Die Wärmebewegung ist dann zwar stark genug, um die orientierte periodische Anordnung in einem Kristallgitter (Abb. 6.3.2-5a) aufzuheben, reicht aber noch nicht aus, um die Wechselwirkungskräfte senkrecht zu den Moleküllängsachsen vollständig zu überwinden.

Nach den vorherrschenden *strukturellen Ordnungsprinzipien* unterscheidet man drei verschiedene Arten kristallin-flüssiger Phasen. In *smektischen Phasen* sind die Moleküle in gut ausgebildeten Schichten parallel gelagert (Abb. 6.3.2-5b). Innerhalb der gegeneinander leicht verschiebbaren Schichten von der Dicke etwa einer Moleküllänge sind sie ohne Periodizität angeordnet und können leicht ihre Plätze wechseln. Auch in den *nematischen Phasen* sind alle Molekülachsen parallelorientiert; sie können sich jedoch in dieser Parallellage in allen drei Raumrichtungen frei gegeneinander bewegen, so daß es nicht zur Ausbildung von Schichten kommt (Abb. 6.3.2-5c). Eng verwandt mit dieser Struktur ist der Aufbau der *cholesterischen Phasen*. Sie bestehen aus 0,2 bis 20 µm dicken Schichten mit nematischer Struktur, in denen auch infolge eines speziellen Aufbaus der betreffenden Moleküle die Vorzugsrichtungen der Molekülachsen von Schicht zu Schicht um einen bestimmten Betrag gleichsinnig verdreht sind. Dadurch entsteht über mehrere solcher Schichten (in Abb. 6.3.2-5c also senkrecht zur Zeichenebene) eine Helixstruktur, d. h. eine schraubenförmige Anordnung der Moleküllängsachsen. Manche der Verbindungen mit kristallin-flüssigen Zustandsbereichen bilden in Abhängigkeit von der Temperatur nacheinander mehrere smektische Phasen aus; jeder Stoff hat dagegen nur eine nematische oder cholesterische Phase, deren Existenzbereich nach höheren Temperaturen durch den Übergang zur isotropen Flüssigkeit (Abb. 6.3.2-5d) im *Klärpunkt* begrenzt ist.

Die ausgeprägte Orientierung in allen kristallin-flüssigen Phasen hat starke *Anisotropien der Eigenschaften*, z. B. der Leitfähigkeit, der Dielektrizitätskonstanten oder des Brechungsindex, zur Folge. Auf ihrer gezielten Ausnutzung beruhen alle seit etwa 1965 in ständig steigendem Umfang eingeführten technischen Anwendungen. Die größte Bedeutung hat heute die Nutzung *elektrooptischer Effekte* mit nematischen und cholesterischen Phasen als *Flüssigkristallanzeigen* [LCD; liquid crystal display], vor allem in Armbanduhren, Taschenrechnern, digitalen Meßgeräten usw. In sehr dünnen Schichten ordnen sich die Moleküle entweder senkrecht, d. h. mit *homöotroper Orientierung*, oder parallel, mit *homogener Orientierung*, oder unter Einfluß eines elektrischen Felds schräg zu den Begrenzungsflächen

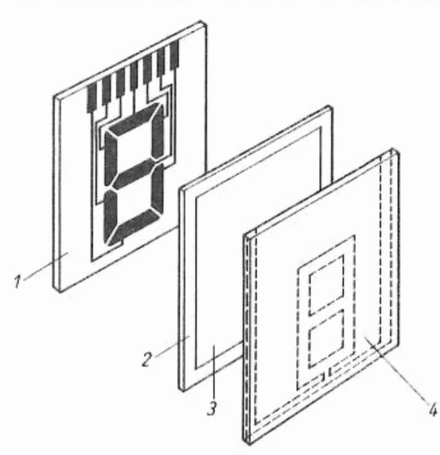


Abb. 6.3.2-6 Schematischer Aufbau einer digitalen 7-Segment-Flüssigkristallanzeige; *1* Glasgrundplatte mit elektrisch leitendem Elektrodenmuster (Segmente mit Anschlüssen), *2* 6 bis 30 µm starker Abstandsrahmen, *3* Flüssigkristallschicht, *4* Deckplatte mit transparentem Gegenelektrodenmuster

an. Ein- oder Ausschalten des Felds bewirkt infolge der Anisotropie der Dielektrizitätskonstanten stets einen Wechsel zwischen diesen Orientierungen, die aufgrund einer außergewöhnlich großen Doppelbrechung deutlich unterschiedliche optische Eigenschaften besitzen. Je nach der Art der kristallin-flüssigen Stoffe und der Art und Stärke der Ansteuerung können eine größere Zahl spezieller Effekte genutzt und damit die Anzeigesysteme unterschiedlichen Anforderungen angepaßt bzw. in unterschiedlicher Weise bis hin zu mehrfarbigen Anzeigen realisiert werden.

Flüssigkristallanzeigen haben einen extrem niedrigen Energieverbrauch, der in der Größenordnung von 0,001 bis 0,5 $\mathrm{mWcm^{-2}}$ liegt, benötigen nur niedrige Betriebsspannungen von etwa 3 bis 20 V, besitzen ein gutes Kontrastverhältnis, das unabhängig von der Umgebungshelligkeit ist, und können im Prinzip beliebig großflächig aufgebaut werden. Für digitale Anzeigen werden vor allem die 7-Segment-Anzeigen für Ziffern (Abb. 6.3.2-6) und die 7 × 5-Punkte-Anzeigen für alphanumerische Zeichen, z. B. Buchstaben, ange-

wandt. Als transparente, elektrisch leitende Deckschichten dienen meist dotierte Zinn- oder Indiumoxidschichten. Hauptnachteile der Flüssigkristallanzeigen sind der begrenzte Betriebstemperaturbereich entsprechend dem kristallin-flüssigen Zustandsbereich der verwendeten Stoffe bzw. Stoffmischungen, heute etwa -20 bis $+60$ °C, die relativ große Empfindlichkeit der Verbindungen gegen stoffliche Einflüsse aus der Umgebung, weswegen sie möglichst hermetisch gegen die Außenwelt abgekapselt werden müssen, sowie für schnell veränderliche Anzeigen die relativ langen Ansprech- bzw. Abklingzeiten in der Größenordnung von 20 bis 300 ms. Ein weiteres bedeutsames Anwendungsgebiet kristalliner Flüssigkeiten mit cholesterischen Phasen ist die *Thermotopographie.* Durch Einwirkung von Wärmeenergie verändert sich der Verdrehungswinkel der einzelnen nematischen Schichten und damit die Helixganghöhe. Dies führt zu Änderungen der Reflexionswellenlänge von auffallendem Licht. Cholesterische Flüssigkeiten, in dünner Schicht bzw. in Form verkapselter Folien auf beliebige Unterlagen aufgetragen, haben daher die Eigenschaft, im weißen Tages- oder bei Kunstlicht innerhalb eines für jede Verbindung charakteristischen Temperaturbereichs das gesamte sichtbare Farbspektrum von rot bis violett zu durchlaufen. Dieser Vorgang ist reversibel, weshalb sich solche Schichten zur kontakthermometrischen Messung von Temperaturen, Temperaturdifferenzen (bis 0,01 K) oder Temperaturänderungen auf Oberflächen eignen. Selbst dynamische Vorgänge wie Wärmeausbreitungsprozesse lassen sich mit Ansprechzeiten von etwa 50 ms visuell oder durch mikroskopische Beobachtungen sehr kleiner Oberflächenbereiche genau verfolgen. Die Beobachtung derartiger stationärer oder dynamischer *Wärmebilder* hat breite Anwendung in der Technik gefunden, z. B. zur zerstörungsfreien Werkstoffprüfung und zur Prüfung der Temperaturverteilung auf elektronischen Bauelementen und Schaltkreisen im Betrieb. Daneben nutzt man die Temperaturabhängigkeit der Reflexionsfarben auch in *thermischen Bildwandlern* aus. Als Strahlungsdetektoren für Mikro-, IR-, UV- oder auch Ultraschallwellen zeigen sie die Erwärmung geeigneter Absorber beim Auftreffen entsprechender Strahlung und die Intensitätsverteilung in solchen Strahlungsfeldern an. Schließlich hat die Flüssigkristalltemperaturanzeige auch bereits *medizinische Anwendungen* erfahren, z. B. für Hauttemperaturmessungen, zum Nachweis von Entzündungen, bei der Krebsdiagnose und Hauttransplantationen. Kristallin-flüssige Systeme können in bestimmten Konzentrationsbereichen auch aus Peptiden oder ähnlichen Verbindungen und Wasser als Lösungsmittel entstehen. Dieser *lyotrope kristallin-flüssige Zustand* hat herausragende Bedeutung für den Stoffwechsel in lebenden Zellen und ist Träger wichtiger Funktionen für den Ablauf der Leitungsvorgänge in den Nerven und im Gehirn (Tafeln 68/69).

6.3.2.2. Thermische und kalorische Eigenschaften. Die *thermischen Ausdehnungskoeffizienten* γ der Flüssigkeiten liegen in der Größenordnung von $10^{-3}\ \mathrm{K}^{-1}$, die *Kompressibilitätskoeffizienten* χ um $10^{-9}\ \mathrm{Pa}^{-1}$ ($10^{-14}\ \mathrm{bar}^{-1}$). Sie werden bei jeder Flüssigkeit in spezifischer Weise von den Wechselwirkungskräften und den Nahordnungszuständen bestimmt. Daher ist es nicht möglich, eine allgemeingültige *thermische Zustandsgleichung* $v = f(m, p, T)$ aufzustellen. In begrenzten Temperatur- und Druckbereichen gilt die Beziehung
$v = v_0[1 + \gamma(T - T_0)]\,[1 - \chi(p - p_0)]$ mit konstanten Werten von γ und χ. v_0 ist dabei das zu dem gewählten Bezugszustand mit der Temperatur T_0 und dem Druck p_0 gehörende Volumen.

Die *innere Energie* der Flüssigkeiten wird wesentlich von der Stärke der jeweiligen Wechselwirkungskräfte zwischen den Teilchen bestimmt. Die Bewegungen der Atome oder Moleküle – Rotationen, Molekülschwingungen und auch Translationen, letztere in Form unregelmäßiger, stark anharmonischer Schwingungen, die zu ständigen geringfügigen Verschiebungen der Teilchenschwerpunkte führen – lassen sich nicht mehr wie bei den Gasen statistisch beschreiben (↑ 6.2.4.), weswegen auch eine allgemeingültige Theorie der *Molwärmen* und ihrer Temperaturabhängigkeit nicht existiert. In den meisten Fällen sind die Temperaturkoeffizienten positiv. Die Differenz $C_p - C_v = (V_0\gamma^2/\chi)\,T$ steigt im allgemeinen mit zunehmender Temperatur stark an.

6.3.2.3. System Flüssigkeit – Dampf. Wie in Gasen läßt sich auch den Teilchen in einer Flüssigkeit nur eine *mittlere Energie* zuschreiben, die durch die Wechselwirkungskräfte und die Temperatur bestimmt wird. Bei jeder Temperatur hat ein Bruchteil aller Teilchen in der an eine Gasphase oder ein Vakuum angrenzenden Phasengrenzschicht einen hinreichenden Energieüberschuß zur Überwindung der zwischenmolekularen Wechselwirkungskräfte und damit zum Ausbrechen aus dem Flüssigkeitsverband und Übertritt in die Nachbarphase. In offenen Systemen bezeichnet man diesen Vorgang als *Verdunstung.* Befinden sich beide Phasen in einem geschlossenen System, so treffen die Teilchen in der Dampfphase bei ihren regellosen Bewegungen auch auf die Gefäßwandungen und üben dort einen dem Gasdruck völlig analogen *Dampfdruck* aus, dessen Größe von der Konzentration der im Gasraum befindlichen Teilchen abhängt. Aus dem Gas-

raum treffen Teilchen auch auf die Flüssigkeitsoberfläche auf, wo sie z. T. reflektiert, hauptsächlich aber von den Flüssigkeitsteilchen so festgehalten werden, daß sie wieder in die flüssige Phase übergehen.

6.3.2.3.1. Im *Verdampfungsgleichgewicht* treten fortwährend in der Zeiteinheit ebenso viele Teilchen aus der Flüssigkeit in den Dampfraum über wie umgekehrt. Dieser Zustand läßt sich nach der allgemeinen thermodynamischen Gleichgewichtsbedingung für Phasengleichgewichte (↑ 6.1.2.4.) durch $p_{s,T}$ = const. beschreiben. Jede Flüssigkeit hat bei gegebener Temperatur einen ganz bestimmten Dampfdruck, den *Sättigungsdampfdruck* p_s. Da dieser bei der Siedetemperatur der Flüssigkeit dem äußeren Druck gerade gleich wird, ist die *Siedetemperatur* bei Normaldruck (1,01325 bar $\approx 10^5$ Pa) eine für jeden Stoff charakteristische Größe. Vom Volumen oder von der Masse der Flüssigkeit ist der Dampfdruck (Tab. 6.3.2-7) unabhängig. Aus der molekulartheoretischen Vorstellung über das Zustandekommen des Dampfdrucks folgt unmittelbar, daß dieser mit steigender Temperatur grundsätzlich zunimmt (Abb. 6.3.2-8). Für die *Temperaturabhängigkeit des Dampfdrucks* ergibt sich aus $\mathrm{d}p_s/\mathrm{d}T = \Delta_V H/[T(V_g - V_l)]$. Hierbei sind V_g bzw. V_l die molaren Volumina des Stoffs im gasförmigen und im flüssigen Zustand. Unter Berücksichtigung von $V_g \gg V_l$ und mit Annahme der Gültigkeit des idealen Gasgesetzes für die Gasphase folgt hieraus die *Clausius-Clapeyronsche-Gleichung* $\mathrm{d}\ln p_s/\mathrm{d}T = \Delta_V H/(RT^2)$ (R molare Gaskonstante). Diese Gleichung gibt das Dampfdruckverhalten vieler Flüssigkeiten bis zu Temperaturen um 100 °C sowie Dampfdrücken bis 10^5 Pa (1 bar) für praktische Zwecke gut wieder. Für größere Temperaturbereiche muß in erster Linie die Temperaturabhängigkeit der – stets positiven – *Verdampfungsenthalpien* $\Delta_V H$ berücksichtigt werden. Diese nehmen mit steigender Temperatur ab; am kritischen Punkt werden sie Null. Die stoffeigenen Verdampfungsenthalpien können mit Hilfe der Clausius-Clapeyronschen oder anderer Dampfdruckgleichungen aus Dampfdruckmessungen ermittelt werden. Sie spiegeln neben ihrer Abhängigkeit von der Molekülgröße deutlich die unterschiedlichen Stärken der Wechselwirkungskräfte zwischen den Flüssigkeitsteilchen wider (↑ Tab. 6.3.2-7).

Tab. 6.3.2-7 Sättigungsdampfdrücke p_s in mbar bei 20 °C, Siedepunkte T_s in °C bei Normaldruck und molare Verdampfungsenthalpien $\Delta_V H$ in kJmol^{-1} bei der Siedetemperatur von verschiedenen Flüssigkeiten

Stoff	p_s	T_s	$\Delta_V H$
H_2O	23,4	100,0	40,6
CH_3OH	129,3	64,7	35,3
C_2H_5OH	58,7	78,4	38,6
$C_2H_5OC_2H_5$	589,2	34,6	26,7
CH_3COCH_3	246,6	56,5	29,2
CCl_4	121,3	76,7	30,0
$CHCl_3$	212,7	61,3	29,3
C_6H_6	100,2	80,1	30,8
$C_6H_5CH_3$	29,3	110,6	32,0
n-C_8H_{18}	13,3	125,6	34,4

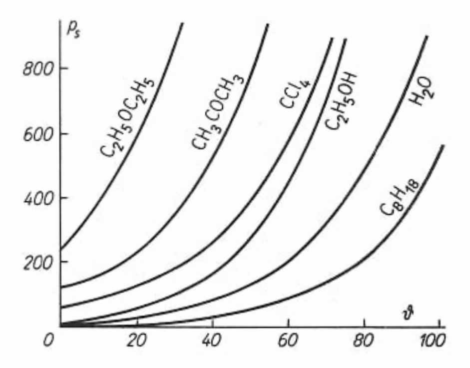


Abb. 6.3.2-8 Temperaturabhängigkeit des Sättigungsdampfdrucks p_s verschiedener Flüssigkeiten; p_s in 10^2 Pa, ϑ in °C

6.3.2.3.2. *Verdampfungs- und Kondensationsgeschwindigkeit.* Die Verdampfungsgeschwindigkeit bei konstanter Temperatur hängt vor allem von der Natur der Flüssigkeit und von der Wärmezufuhr ab. Erfolgt die Verdampfung in ein Vakuum und werden alle aus der Flüssigkeitsoberfläche ausgetretenen Teilchen sofort von ihr entfernt, dann ist die Anzahl $\dot{N}$ der pro cm² in 1 s verdampfenden Teilchen nach Hertz und Knudsen $\dot{N} = N_A p_s/\sqrt{2\pi MRT}$ (N_A Avogadrosche Zahl, p_s Sättigungsdampfdruck bei der Temperatur T, M molare Masse, R molare Gaskonstante). Diese *ideale Verdampfung* hat als *Molekulardestillation* Bedeutung für die Trennung chemisch ähnlicher Stoffe, da sich die Verdampfungsgeschwindigkeiten oft stärker unterscheiden als die Dampfdrücke.

Bei der *Verdampfung in eine Gasatmosphäre* wird durch die Vielzahl der Zusammenstöße im Dampfraum stets ein Teil der verdampften Teilchen auf die Flüssigkeitsoberfläche zurückgelenkt und von der flüssigen Phase wieder aufgenommen. $\dot{N}$ ist dann immer kleiner als bei der Verdampfung in ein Vakuum und hängt von der Geschwindigkeit und vom Umfang ab, mit der die aus der Oberfläche der Flüssigkeit gerade ausgetretenen Teilchen durch Diffusion

und Konvektion aus ihrer Nähe wegtransportiert werden.

Beim *Sieden* gehen nicht nur an der Phasengrenze Teilchen der Flüssigkeit in den Dampfraum über, sondern es entstehen auch im Inneren der flüssigen Phase *Dampfblasen*. Der geschwindigkeitsbestimmende Teilvorgang ist dabei die *Kavitation*, d. h. das Entstehen von winzigen *Gasbläschenkeimen* durch „Auseinanderreißen" der Flüssigkeit. Hierbei müssen die zwischenmolekularen Wechselwirkungskräfte und der Krümmungsdruck (↑ 6.6.2.3.2.) überwunden werden, der bei den sehr kleinen Bläschen infolge der starken Oberflächenkrümmung hohe Werte annimmt. Nach Abschätzungen liegt der Gegendruck „Kavitationslöcher" in der Größenordnung von 10^9 Pa (10^4 bar). Um diese *Keimbildungshemmungen* zu überwinden, sind bei sehr reinen Flüssigkeiten, in denen die Gasblasenkeime nicht z. B. an winzigen Schwebstoffpartikeln gebildet werden können, oftmals beträchtliche *Überhitzungen* nötig. Bei Ether hat man z. B. einen „Siedeverzug" um über 100 K beobachten können.

Analoge Hemmungen treten auch bei dem der Verdampfung entgegengesetzten Vorgang, der *Kondensation* auf. Wenn eine flüssige Phase noch nicht vorhanden ist, muß die normale Kondensationstemperatur, bei Wasserdampf 100 °C, stets unterschritten werden, damit sich *Tröpfchenkeime* bilden können. Der Dampf befindet sich dann in einem metastabilen Zustand. Es treten Keimbildungshemmungen dadurch auf, daß der Dampfdruck sehr kleiner Flüssigkeitströpfchen durch den Krümmungsdruck um Größenordnungen ansteigt. Die Wahrscheinlichkeit dafür, daß beim zufälligen Zusammenstoß mehrerer Teilchen ein Keim wenigstens für eine gewisse Zeit stabil bleibt und weiter wächst, ist in reinsten Dämpfen selbst bei hohen Übersättigungsgraden gering. Sie steigt aber schon bei geringen Übersättigungen stark an, wenn im Dampfraum Fremdpartikeln vorhanden sind, an denen die Kondensation erfolgen kann. Wasserdampf in der Luft kondensiert z. B. vorwiegend an Staubpartikeln. Zur Vermeidung von Hagel versucht man in entsprechende Wolken solche festen Stoffe in sehr feiner Verteilung zu injizieren, an deren Oberflächen sich infolge ihrer speziellen stofflichen, strukturellen und energetischen Eigenschaften H_2O-Moleküle möglichst günstig anlagern und damit kondensieren können. Durch eine ähnliche Wirkung entstehen in Höhen um 10 km die Kondensstreifen hinter den Triebwerken von Düsenflugzeugen. Besonders stark keimfördernd sind elektrisch geladene Teilchen, was z. B. in der Wilsonschen Nebelkammer ausgenutzt wird (↑ 3.2.1.2.2.).

Bei Anwesenheit der flüssigen Phase hängt die *Kondensationsgeschwindigkeit* neben der Abführung der freiwerdenden Kondensationsenthalpie $\Delta_K H$ ($= -\Delta_V H$) vor allem vom Antransport der Teilchen im Dampfraum an die Flüssigkeitsoberfläche ab.

6.3.3. Bildung und Zustandseigenschaften flüssiger Mischphasen

Die Bildung *flüssiger Stoffmischungen* aus den reinen Stoffen wird in erster Linie von den im allgemeinen unterschiedlichen *zwischenmolekularen Wechselwirkungskräften* und *Ordnungszuständen* in ihnen bestimmt. Beim Zusammenbringen verschiedener Flüssigkeiten bzw. einer Flüssigkeit mit gasförmigen oder festen Stoffen treten deshalb Veränderungen der Eigenschaften auf.

Besonders augenfällig ist bereits die Tatsache, daß sich zwei Flüssigkeiten entweder unbegrenzt oder nur begrenzt oder praktisch überhaupt nicht miteinander vermischen bzw. daß bestimmte gasförmige oder feste Stoffe sich in einer Flüssigkeit lösen, andere dagegen praktisch nicht oder nur in beschränktem Umfang. Sehr große Unterschiede gegenüber den reinen Flüssigkeiten zeigen solche flüssigen Mischphasen, in denen elektrisch geladene Teilchen, Ionen, enthalten sind (↑ 6.3.4.).

6.3.3.1. Ursachen und Erscheinungsformen der Mischbarkeit. Hauptursache dafür, daß sich verschiedene Stoffe vermischen, ist die *thermische Molekularbewegung* und damit die Tendenz aller Teilchen, sich im ganzen zur Verfügung stehenden Volumen durch Diffusion auszubreiten. Thermodynamisch entspricht der Unterschied zwischen den Ordnungszuständen in den reinen Komponenten und in der Mischphase dem Auftreten einer stets positiven *Mischungsentropie* $\Delta_M s$.

Neben der Teilchenbewegung spielen aber auch die *Wechselwirkungskräfte* zwischen den verschiedenen Teilchenarten eine entscheidende Rolle. Ihr Wirken kommt thermodynamisch noch zusätzlich durch eine positive oder negative *Mischungsenthalpie* $\Delta_M h$ zum Ausdruck. Grundsätzlich ist eine Mischungs- oder Lösungsbildung nur möglich, wenn die freie Enthalpie der Mischphase unter den gegebenen Bedingungen kleiner wird als die Summe der freien Enthalpien der reinen Komponenten, d. h. wenn nach $\Delta_M g = \Delta_M h - T \Delta_M s$ die *freie Mischungsenthalpie* $\Delta_M g$ negativ ist (↑ 6.1.2.3.2.). Andernfalls bilden die verschiedenen Komponenten ein heterogenes Mehrphasensystem.

6.3.3.1.1. *Systematisierung.* Die zwischenmolekularen Wechselwirkungskräfte können die Bildung einer Mischphase sowohl begünstigen als auch hemmen bzw. verhindern. Dies hängt davon ab, ob die Kräfte F_{AB} zwischen den

unterschiedlichen Teilchen, z. B. den Lösungsmittelmolekülen A und den Teilchen des gelösten Stoffs B, größer, etwa gleich oder kleiner als die Mittelwerte der entsprechenden Kräfte F_{AA} und F_{BB} zwischen den Teilchen ein und derselben Art sind. Damit läßt sich folgende Einteilung der Mischphasen vornehmen:

1. $F_{AB} > \frac{1}{2}(F_{AA} + F_{BB})$. Die Vermischung wird begünstigt; jedes Teilchen versucht, sich möglichst weitgehend mit Teilchen der anderen Art zu umgeben. Dabei wird Energie frei, $\Delta_M h < 0$; außerdem bewirkt die stärkere Wechselwirkung zwischen den Teilchen A und B häufig eine Volumenkontraktion, $\Delta_M v < 0$, d. h., das Volumen der Mischphase ist kleiner als die Summe der Volumina der Komponenten vor der Mischung. Beispiele dafür sind Mischungen aus Wasser und Alkoholen.
2. $F_{AB} \approx \frac{1}{2}(F_{AA} + F_{BB})$. In diesem Fall erfahren die Wechselwirkungskräfte praktisch keine Veränderungen. Daher tritt auch keine Enthalpieänderung und keine Volumenänderung auf, $\Delta_M h \approx 0$ und $\Delta_M v \approx 0$. Man nennt solche Mischphasen *ideale Mischungen* (↑ 6.3.3.3.1.).
3. $F_{AB} < \frac{1}{2}(F_{AA} + F_{BB})$. Die Wechselwirkungskräfte wirken dem Vermischungsbestreben entgegen. Falls überhaupt eine Mischphasenbildung eintritt, sind $\Delta_M h$ und $\Delta_M v > 0$. Zu diesem Mischungstyp gehören z. B. Mischungen aus Benzen und Alkohol. Bei der ungeheuren Vielfalt der Teilchenarten und ihrer Wechselwirkungen sowie der möglichen Ordnungszustände ist es im allgemeinen nicht möglich, mit Sicherheit vorherzusagen, wie sich die Komponenten zueinander verhalten werden. Außerdem können infolge spezieller Teilcheneigenschaften auch Abweichungen von der getroffenen Einteilung vorkommen.

6.3.3.1.2. Die *Mischungsenthalpien* und *Mischungsentropien* und damit die Eigenschaften der Mischphasen hängen von der Art der Komponenten und den Zustandsbedingungen der Mischung, vor allem von den Mengenverhältnissen der Komponenten (↑ 6.1.2.1.1.) sowie von der Temperatur ab.

Bei binären Mischungen durchlaufen die *Mischungsenthalpien* in Abhängigkeit vom Molenbruch häufig parabolische Kurven mit Extremwerten in der Nähe von $x = 0{,}5$ (Abb. 6.3.3-1). Bedingt durch spezielle Strukturverhältnisse in den reinen Komponenten können jedoch auch stark abweichende Verläufe auftreten (↑ Abb. 6.3.3-1, Kurve *3*).

Bei idealen Mischungen gilt für die Mischungsentropie $\Delta_M s = -R \sum x_i \ln x_i$ (R molare Gaskonstante, x_i Molenbruch der Komponente i in der Mischphase); sie ist im Unterschied zur Mischungsenthalpie *stets* verschieden von Null. Ideales Verhalten wird beobachtet, wenn sich die Komponenten in ihren physikalischen und chemischen Eigenschaften nur wenig voneinander unterscheiden, z. B. bei Hexan/Heptan, *o*-Xylol/*p*-Xylol oder Chlorbenzen/Brombenzen. Größere Veränderungen in den Ordnungszuständen führen dagegen zu größeren oder kleineren Werten von $\Delta_M s$. Besonders große Werte ergeben sich, wenn die Teilchen in den reinen Komponenten in einer ausgeprägten Struktur vorliegen, die beim Mischen durch die anderen Teilchen weitgehend aufgehoben wird. Kleinere Werte sind demgegenüber stets zu erwarten, wenn die unterschiedlichen Teilchen infolge besonders starker Wechselwirkungen in der Mischphase stärker zueinander geordnet sind oder eine in den reinen Komponenten bereits vorhandene Ordnung noch verstärkt wird. Alle derartigen Veränderungen wirken sich auch auf die Volumina der Mischphasen aus (↑ 6.1.2.1.5.).

6.3.3.1.3. *Vollständige und teilweise Mischbarkeit.* Bei den beiden Mischungstypen 1 und 2 (↑ 6.3.3.1.1.) wird man im allgemeinen mit einer *unbegrenzten* oder *vollständigen Mischbarkeit* im ganzen Temperaturbereich des flüssigen Zustands rechnen können. Beim Typ 3 hängt es dagegen von den jeweiligen Werten von $\Delta_M h$ und $\Delta_M s$ ab, inwieweit eine Mischphasenbildung vor sich geht. Dann kann sich die Mischbarkeit verschiedener Stoffe merklich mit der Temperatur ändern oder auf bestimmte Temperatur- und Konzentrationsbereiche beschränken.

Wenn in solchen Systemen die Molekularbewegung nicht mehr ausreicht, um den Einfluß der

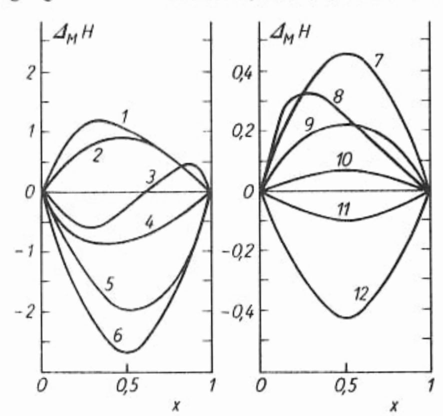


Abb. 6.3.3-1 Molare Mischungsenthalpien $\Delta_M H$ in $kJmol^{-1}$ verschiedener Flüssigkeiten bei 25 °C in Abhängigkeit vom Molenbruch x jeweils der zweiten Komponenten;
1 C_6H_6/C_2H_5OH, *2* Cyclo-C_6H_{12}/C_6H_6,
3 $CH_3OH/CHCl_3$, *4* H_2O/CH_3OH,
5 $CH_3COCH_3/CHCl_3$, *6* $C_2H_5OC_2H_5/CHCl_3$,
7 n-$C_6H_{14}/C_6H_5CH_3$, *8* CCl_4/CH_3COCH_3,
9 $CCl_4/CHCl_3$, *10* $C_6H_6/C_6H_5CH_3$,
11 $C_6H_5CH_3/C_6H_5Cl$, *12* $C_6H_6/CHCl_3$

zwischenmolekularen Wechselwirkungskräfte zu kompensieren, kommt es bei einer für jedes System charakteristischen *kritischen Mischungs-* oder *Lösungstemperatur* zu einer spontanen *Entmischung*. Die ursprünglich homogene Mischphase zerfällt dann in zwei oder auch mehrere *konjugierte Phasen* unterschiedlicher Zusammensetzung, die durch scharfe Phasengrenzen voneinander getrennt sind und miteinander im thermodynamischen Gleichgewicht stehen. Diese koexistierenden Phasen sind bei den betreffenden Temperaturen gesättigte Lösungen der einen Komponente in der anderen. Man unterscheidet drei Arten derartiger Systeme.

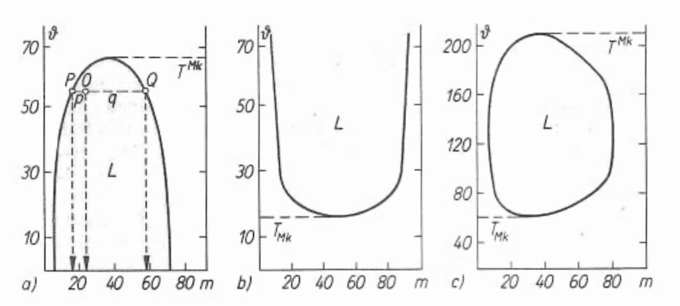


Abb. 6.3.3-2 Temperatur-Konzentrations-Mischungsdiagramm; a) System Wasser/Phenol (m Masse-% Phenol), b) System Wasser/Triethylamin (m Masse-% Triethylamin), c) System Wasser/Nikotin (m Masse-% Nikotin); ϑ in °C

Bei Systemen mit einer *oberen kritischen Mischungstemperatur* T^{Mk} (Abb. 6.3.3-2a) liegen innerhalb einer *Mischungslücke L*, z. B. bei der Gesamtzusammensetzung O, die Komponenten in getrennten, gegenseitig gesättigten Mischphasen mit den Zusammensetzungen P und Q vor. Ihr Mengenverhältnis ist gegeben durch $m_P/m_Q = q/p$. Erst bei Temperaturen oberhalb T^{Mk} bilden die beiden Komponenten im ganzen Konzentrationsbereich eine einzige homogene Mischphase.

Besondere Verhältnisse, z. B. echte chemische Verbindungsbildung bei niedrigen Temperaturen, sind die Ursache für die Existenz von Systemen mit einer *unteren kritischen Mischungstemperatur* T_{Mk} (Abb. 6.3.3-2b). Hier bilden die Komponenten unterhalb dieser Temperatur eine homogene Mischphase, mit steigender Temperatur tritt Entmischung ein, manchmal bis zur völligen Trennung der Komponenten.

Schließlich gibt es noch Systeme mit einer unteren *und* einer oberen kritischen Mischungstemperatur (Abb. 6.3.3-2c). Sie haben dann eine *geschlossene Mischungslücke L*, die Komponenten sind nur bei Temperaturen unterhalb und oberhalb dieser vollständig miteinander mischbar.

6.3.3.1.4. *Lösungen von Gasen in Flüssigkeiten.* Nach dem *Henry-Daltonschen Gesetz* (1807) ist der Molenbruch x_G eines gelösten Gases G in der Lösung seinem Partialbruch p_G über der flüssigen Mischphase proportional, $x_G = A_G p_G$. Der stoffeigene Proportionalitätsfaktor A_G ist der *Henrysche Löslichkeits-* oder *Absorptionskoeffizient*, der experimentell bestimmt werden muß und von der Temperatur abhängt. Er gibt die Konzentration des Gases in der Lösung an, wenn sein Partialdruck über ihr $1{,}013 \cdot 10^5$ Pa (1,013 bar) beträgt.

Das Henry-Daltonsche Gesetz ist ein Grenzgesetz für geringe Löslichkeiten und ideales Verhalten der Gasphase. Trotzdem ist sein Gültigkeitsbereich recht beträchtlich. Es gilt jedoch nur für solche Gase, die sich in den betreffenden flüssigen Lösungsmitteln lösen, ohne mit dessen Teilchen in chemische Reaktion zu treten.

Für praktische Zwecke benutzt man anstelle des Henryschen häufiger den *technischen Löslichkeitskoeffizienten*. Dieser gibt das Volumen des Gases in cm³ an, das die bei der Temperatur T und einem Druck von $0{,}981 \cdot 10^5$ Pa (0,981 bar) des Gases in 1 g der Flüssigkeit gelöste Gasmenge unter Normalbedingungen einnehmen würde. In Wasser beträgt er z. B. bei 20 °C für O_2 $3{,}06 \cdot 10^{-7}\,cm^3g^{-1}Pa^{-1}$ und N_2 $1{,}55 \cdot 10^{-7}\,cm^3g^{-1}Pa^{-1}$. Die Löslichkeit der permanenten und vieler anderer tiefsiedender Gase in Flüssigkeiten ist im allgemeinen nicht sehr groß; gelöste Salze setzen die Löslichkeit meist noch stark herab, und zwar für alle Gase etwa im gleichen Verhältnis. Mit steigender Temperatur nimmt die Löslichkeit stets ab; deshalb läßt sich z. B. Wasser durch Kochen weitgehend von gelöster Luft befreien. Das angeführte Beispiel hat große Bedeutung für das Leben im Wasser. Wie die Zahlenwerte zeigen, ist das Verhältnis der Konzentrationen $O_2 : N_2$ in der Lösung merklich größer als in der atmosphärischen Luft, der Sauerstoffanteil entspricht etwa 36 Volumenprozent im Vergleich zu rund 21 Volumenprozent in der Luft.

6.3.3.2. Auflösung fester Stoffe in Flüssigkeiten

6.3.3.2.1. *Lösungsenthalpien und Sättigungskonzentrationen.* Bei der Bildung flüssiger Mischphasen aus einem flüssigen Lösungsmittel und

festen, kristallinen Stoffen gehen in die *Lösungsenthalpien* zwei Anteile ein. Zum einen muß Energie aufgewandt werden, um die Bausteine des kristallinen Stoffs in den flüssigen Zustand zu überführen. Diese zur Überwindung der Gitterkräfte notwendige Energie entspricht der Schmelzenthalpie. Zum zweiten tritt die eigentliche Mischungsenthalpie auf, die je nach der Art und Stärke der Wechselwirkungskräfte in der Mischphase, d. h. je nach der Zuordnung des Systems zum Mischungstyp 1, 2 oder 3 (↑ 6.3.3.1.1.) unterschiedliche Größe und Vorzeichen haben kann. Insgesamt tritt als Lösungsenthalpie nur die Differenz zwischen den beiden Anteilen in Erscheinung, die deshalb positiv oder negativ, in selteneren Fällen auch praktisch Null sein kann (Tab. 6.3.3-3) und die außerdem meist stark vom Mengenverhältnis der beiden Komponenten in der Lösung sowie von der Temperatur abhängt.

Die in 6.3.3.1.3. begründete Beschränkung der Mischbarkeit auf bestimmte Temperatur- und Konzentrationsbereiche für Systeme aus mehreren Flüssigkeiten findet sich auch bei Lösungen fester Stoffe, jedoch ist hier die Löslichkeit häufig auch dann begrenzt, wenn ein Mischungstyp 1 vorliegt. Ursache dafür sind in erster Linie die mit den starken Wechselwirkungen F_{AB} einhergehenden besonders großen Änderungen der Flüssigkeitsstruktur. Viele feste Stoffe lösen sich daher nur bis zu einer bestimmten *Sättigungskonzentration* c_s (↑ Tab. 6.3.3-3), die nur von der Temperatur, nicht aber vom Mengenverhältnis der reinen Stoffe abhängt. Ihre Temperaturabhängigkeit ist näherungsweise gegeben durch $\mathrm{d}\ln c_s/\mathrm{d}T = \Delta_L H^{\infty}/(RT^2)$ ($\Delta_L H^{\infty}$ molare Lösungsenthalpie bei unendlicher Verdünnung, R molare Gaskonstante). Es hängt dann vom Vorzeichen der Lösungsenthalpie ab, ob c_s mit Temperatursteigerung zu- oder abnimmt. In sehr vielen Fällen wird eine Zunahme beobachtet.

6.3.3.2.2. *Auflösungsgeschwindigkeit.* Bei der Bildung flüssiger Mischphasen aus einer Flüssigkeit und einem Festkörper ist in der Praxis auch die Auflösungsgeschwindigkeit von Bedeutung. Eine Auflösung kann nur von der Oberfläche des festen Stoffs her vor sich gehen. Sie erfolgt daher in zwei nacheinander und mit unterschiedlichen Geschwindigkeiten ablaufenden Teilvorgängen, nämlich dem eigentlichen Übergang der Oberflächenbausteine aus dem Gitterverband in eine unmittelbar an den Festkörper angrenzende Schicht der Flüssigkeit und Diffusionsvorgängen.

In den meisten Fällen ist die Auflösung *transportbestimmt*, d. h. durch die Geschwindigkeit bestimmt, mit der die gelösten Teilchen von der Auflösungsstelle in das Innere der Flüssigkeit bzw. Lösungsmittelmoleküle von dort an die Festkörperoberfläche diffundieren. Unmittelbar an die Oberfläche grenzt eine Flüssigkeitsschicht in der die Sättigungskonzentration herrscht und von der aus die Konzentration des Gelösten bis ins Innere der flüssigen Mischphase absinkt. Die dadurch bewirkte Diffusion kann durch Konvektion der Flüssigkeit, z. B. Umrühren oder Umwälzen, beschleunigt werden. Stets aber bleibt unmittelbar an der Festkörperoberfläche die *Diffusionsschicht* mit einer Dicke von etwa 0,01 bis 0,1 mm erhalten, durch die alle Teilchen nur durch Diffusion hindurchtreten können. Die Auflösungsgeschwindigkeit kann dann näherungsweise durch die *Nernst-Noyes-Whitney-Gleichung*

$$(1) \quad -\mathrm{d}n_B/\mathrm{d}t = (DA/\delta)\,(c_{s,B} - c_B)$$

beschrieben werden. In ihr bedeuten $-\mathrm{d}n_B/\mathrm{d}t$ die Menge des festen Stoffs B, die in der Zeiteinheit aus dem Gitterverband in die Lösung übergeht, D den Diffusionskoeffizienten dieser Teilchen, A die Oberflächengröße des festen Stoffs, δ die Dicke der Diffusionsschicht, $c_{s,B}$ die Sättigungskonzentration des Stoffs B im betreffenden Lösungsmittel und c_B die Konzentration im Innern der Lösung. Die Auflösungsgeschwindigkeit kann erhöht werden durch mechanische Bewegung der Flüssigkeit (Verkleinerung von δ), durch Zerkleinern des Festkörpers (Vergrößerung von A) sowie in geringerem Maße durch Erhöhung der Temperatur (Vergrößerung von D und meist auch von $c_{s,B}$).

6.3.3.3. Ideale und reale Mischungen bzw. Lösungen

6.3.3.3.1. *Allgemeine Kennzeichnung.* Aus der Vielzahl flüssiger Mischphasen hebt sich die

Tab. 6.3.3-3 Molare Lösungsenthalpien $\Delta_L H$ in kJ mol^{-1} bei der Auflösung von 1 mol verschiedener organischer und anorganischer Stoffe in jeweils ≧200 mol Wasser bei 20 °C und Sättigungskonzentrationen c_s in g pro 100 g Wasser bei 20 °C

Stoff	$\Delta_L H$	c_s
Harnstoff	+15,4	108
Rohrzucker	+3,8	204
Milchzucker	+7,7	19
Oxalsäure	+9,6	9,5
Benzoesäure	+27,2	0,3
NaCl	+5,0	35,85
$NaNO_3$	+21,3	88,0
Na_2CO_3	−23,4	21,6
KCl	+18,8	34,35
KNO_3	+36,0	31,5
NH_4Cl	+16,3	37,4
$MgCl_2$	−150,6	54,25
$CuSO_4$	−66,5	20,9
$AgNO_3$	+22,6	215,5

kleine Gruppe der zum Mischungstyp 2 (↑ 6.3.3.1.1.) gehörenden Systeme heraus. Ihr allgemeines Kennzeichen ist, daß sich nicht nur die Enthalpien und die Volumina, sondern mit Ausnahme der Entropie auch alle anderen extensiven (masseproportionalen) Zustandseigenschaften z *additiv* aus den entsprechenden molaren Größen Z und Stoffmengen n der Komponenten A, B ... zusammensetzen: $z_M = n_A Z_A + n_B Z_B + \ldots$ Man bezeichnet solche Systeme als *ideale* oder *vollkommene Mischungen*; die Bezeichnung „ideal" hat hier eine ganz andere Bedeutung als bei „idealen Gasen" oder „idealen Festkörpern" (s. a. 6.2.1. und 6.4.).

Alle anderen, zu den Mischungstypen 1 und 3 gehörenden flüssigen Mischungen heißen *reale Mischphasen*. Ihnen ist gemeinsam, daß die extensiven Eigenschaften *nicht additiv* sind und daß die auftretenden Mischungseffekte von der Natur der jeweiligen Komponenten abhängen. Um diese Eigenschaften dennoch mit einfachen, den für ideale Mischphasen analogen Beziehungen beschreiben zu können, muß man sich der partiellen molaren Größen bedienen (↑ 6.1.2.1.5.).

Sind die Konzentrationen aller Komponenten realer Mischphasen bis auf eine, das Lösungsmittel, sehr klein, dann zeigen diese Lösungen in bestimmten makroskopischen Eigenschaften ein „ideales" Verhalten, wie es bei den idealen Mischphasen über den ganzen Konzentrationsbereich auftritt. Natürlich bestehen auch in diesem *Grenzzustand* nach wie vor starke Wechselwirkungskräfte F_{AA} zwischen den Teilchen des Lösungsmittels und F_{AB} zwischen denen der gelösten Stoffe und des Lösungsmittels. Die Teilchen der gelösten Stoffe beeinflussen sich aber infolge ihrer geringen Konzentration gegenseitig praktisch nicht mehr, $F_{BB}, F_{CC}, \ldots = 0$. Damit werden die Eigenschaften solcher *ideal verdünnten Lösungen* ausschließlich durch die Natur des Lösungsmittels bestimmt. Von den gelösten Stoffen ist dagegen nur die Konzentration von Einfluß, aber nicht mehr ihre Art. Derartige Eigenschaften sind der Dampfdruck und die mit ihm zusammenhängenden Erscheinungen der Dampfdruckerniedrigung, der Siedetemperaturerhöhung, der Gefriertemperaturerniedrigung sowie des osmotischen Drucks. Man faßt diese unter der Bezeichnung *kolligative Eigenschaften* zusammen.

6.3.3.3.2. *Dampfdrücke idealer Mischungen.* Wenn die Wechselwirkungskräfte zwischen den Teilchen einer flüssigen Mischphase praktisch gleich sind, ist es für den Übergang eines Teilchens aus der flüssigen in die Gasphase gleichgültig, ob es nur von gleichartigen (↑ 6.3.2.3.) oder auch von andersartigen Nachbarteilchen umgeben ist. Die Dampfdrücke der verschiedenen Komponenten werden lediglich proportional ihrem Anteil in der Flüssigkeit verringert. Für den Dampfdruck p_M einer idealen Mischung gilt daher, solange die Gasphase den Gesetzen des idealen Gaszustands gehorcht $p_M = p_A + p_B + \ldots = x_A p_A^0 + x_B p_B^0 + \ldots$ $p_A, p_B, \ldots$ sind die Partialdampfdrücke, $x_A, x_B, \ldots$ die Molenbrüche der Komponenten $A, B, \ldots$ in der flüssigen Mischphase und $p_A^0, p_B^0, \ldots$ die Dampfdrücke der reinen Komponenten bei der gleichen Temperatur. Die lineare Änderung der Partialdampfdrücke mit den Molenbrüchen kommt gut in den entsprechenden *Dampfdruckdiagrammen* (Abb. 6.3.3-4) zum Ausdruck.

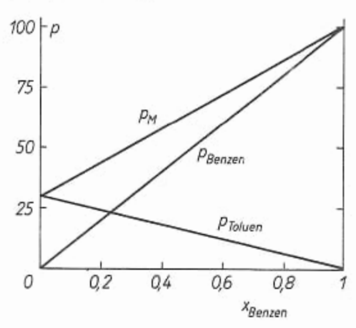


Abb. 6.3.3-4 Partialdampfdrücke und Gesamtdampfdruck p_M flüssiger Mischphasen aus Benzen und Toluen bei 20 °C; p in 10^2 Pa

Insbesondere bei Lösungen fester Stoffe in Flüssigkeiten hat oftmals nur das Lösungsmittel A einen merklichen Dampfdruck. Auch dann gilt für ideal verdünnte Lösungen $p_A = p_M = x_A p_A^0$; der Dampfdruck p_A ist stets kleiner als der Dampfdruck p_A^0 des reinen Lösungsmittels bei der gleichen Temperatur, denn x_A ist immer kleiner als 1. Diese thermodynamisch streng ableitbare Gesetzmäßigkeit ist das *Raoultsche Gesetz der Dampfdruckerniedrigung*. Für ideale binäre Mischphasen läßt es sich mit $x_B = 1 - x_A$ auch mit $x_B = (p_A^0 - p_A)/p_A^0 = \Delta p/p_A^0$ angeben. Die *relative Dampfdruckerniedrigung* $\Delta p/p_A^0$ des Lösungsmittels ist gleich dem Molenbruch x_B des gelösten Stoffs.

6.3.3.3.3. *Dampfdrücke realer Mischungen.* Je nachdem, ob die Wechselwirkungskräfte F_{AB} zwischen den verschiedenartigen Teilchen einer flüssigen Mischphase größer oder kleiner sind als die Mittelwerte der Kräfte F_{AA} und F_{BB}, ergeben sich *negative* oder *positive Abweichungen* vom Raoultschen Gesetz. Es muß dann in der allgemeingültigen Form $p_A = a_A p_A^0$ bzw. $p_B = a_B p_B^0$ mit den Aktivitäten a der Komponenten (↑ 6.1.2.4.1.) anstelle der Molenbrüche x gerechnet werden.

Kleinere Partialdampfdrücke als bei idealem Verhalten werden immer dann auftreten, wenn die Teilchen in der Flüssigkeit infolge stärkerer Wechselwirkungen nur in geringerem Umfang in die Gasphase übertreten können. Wie in Abb. 6.3.3-5 dargestellt, durchläuft der Gesamtdampfdruck p_M dann ein Minimum. Bei ihm haben flüssige Mischphase und Dampfphase exakt die gleiche Zusammensetzung. Bei allen anderen Mischungsverhältnissen ist, wie bei idealen Mischungen im gesamten Mischungsbereich, die flüchtigere Komponente im Dampf angereichert. Allgemein gilt für die Zusammensetzung der Dampfphase $x_A^g = p_A/p_M$ und $x_B^g = p_B/p_M$, wobei x_A^g und x_B^g die Molenbrüche der Komponenten in der Dampfphase sind.

Bei Systemen mit positiven Abweichungen liegen die Kurven der Partialdampfdrücke und des Gesamtdrucks oberhalb der Raoultschen Geraden (Abb. 6.3.3-6). Infolge der schwächeren Wechselwirkungen in der Flüssigkeit können hier die Teilchen in erhöhtem Umfang in die Gasphase übergehen. Der Gesamtdampfdruck zeigt dann bei einer bestimmten Zusammensetzung ein Maximum; die flüssige und die Dampfphase haben dort genau die gleiche Zusammensetzung. Bei allen anderen Zusammensetzungen der flüssigen Phase ist wiederum die flüchtigere Komponente im Dampf angereichert. Wie die Abb. 6.3.3-5 und 6.3.3-6 zeigen, gehen die realen binären Mischphasen im Bereich sehr niedriger Konzentrationen einer der Komponenten mehr und mehr in den Zustand der ideal verdünnten Lösung über.

Flüssige Mischphasen sieden bei einer Temperatur, bei der ihr Gesamtdampfdruck den Wert des äußeren Drucks erreicht. Diese *Siedetemperaturen* sind von der Zusammensetzung der Flüssigkeit abhängig und müssen bei realen Mischungen experimentell bestimmt werden. Für praktische Zwecke sehr nützlich sind die *Siedediagramme*, in denen die Siedetemperaturen über der Zusammensetzung aufgetragen sind. Einem Minimum bzw. einem Maximum im Dampfdruckdiagramm einer Mischung entspricht ein Maximum bzw. ein Minimum im Siedediagramm. Da in diesen Punkten die flüssige Phase und die Dampfphase die gleiche Zusammensetzung haben, geht hier bei einer *Destillation* ein konstant siedendes, *azeotropes Gemisch*, über. Bei allen anderen Zusammensetzungen der Flüssigkeiten ist es wie bei idealen Mischungen dagegen prinzipiell möglich, infolge der unterschiedlichen Zusammensetzungen der beiden Phasen eine der Komponenten durch Destillation, also auf dem Umweg über die Gasphase, teilweise abzutrennen. Je nach der Lage des azeotropen Punkts im Siedediagramm läßt sich diese Trennung bei mehrfachen Wiederholungen, d. h. durch *fraktionierte Destillation*, durch Reduzierung des äußeren Drucks, d. h. *Vakuumdestillation*, oder durch Zugabe eines geeigneten anderen Stoffs, der mit der im Rückstand angereicherten Komponente der ursprünglichen Mischung ein günstigeres azeotropes Gemisch bildet, verbessern. Destillationen spielen im großtechnischen Maßstab in der chemischen Industrie oder bei der Erdölaufbereitung eine außerordentlich wichtige Rolle.

6.3.3.3.4. *Siedepunktserhöhung und Gefrierpunktserniedrigung.* Lösungen, bei denen nur das Lösungsmittel merklich flüchtig ist, weisen stets einen niedrigeren Dampfdruck als das reine Lösungsmittel bei der gleichen Tempera-

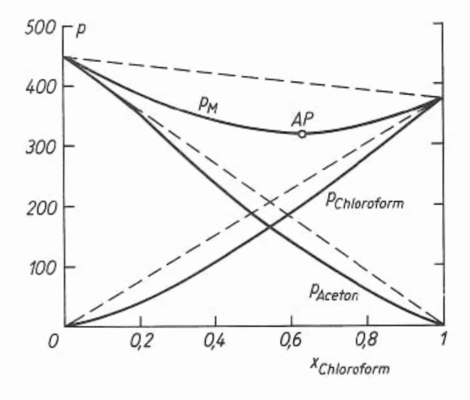


Abb. 6.3.3-5 Dampfdruckdiagramm des Systems Aceton/Chloroform bei 35 °C; p in 10^2 Pa, theoretischer Verlauf bei Gültigkeit des Raoultschen Gesetzes *gestrichelt*, *AP* azeotroper Punkt

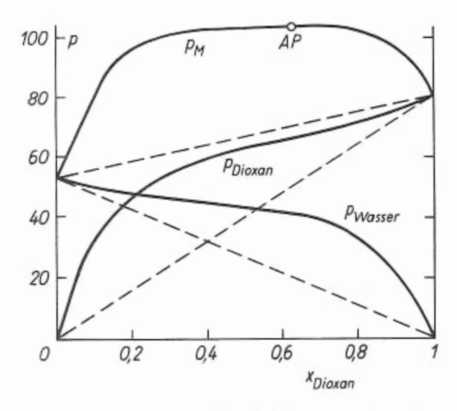


Abb. 6.3.3-6 Dampfdruckdiagramm des Systems Wasser/1,4-Dioxan bei 35 °C; p in 10^2 Pa, theoretischer Verlauf bei Gültigkeit des Raoultschen Gesetzes *gestrichelt*, *AP* azeotroper Punkt

tur (↑ 6.3.3.3.2.) und daher eine *Siedepunktserhöhung* auf (Abb. 6.3.3-7).
Bei idealen Mischungen und ideal verdünnten Lösungen ist ΔT_S nur von der Art der flüchtigen Komponente sowie der Konzentration der nichtflüchtigen Bestandteile abhängig, nicht

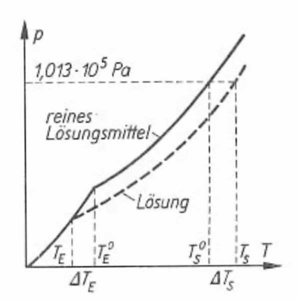


Abb. 6.3.3-7 Siedepunktserhöhung ΔT_S und Gefrierpunktserniedrigung ΔT_E bei flüssigen Mischphasen mit einer flüchtigen Komponente

aber von deren Natur. Für binäre Mischphasen dieser Art ist ΔT_S gegeben durch das *van't-Hoffsche Gesetz der Siedepunktserhöhung*

(2) $$\Delta T_S = T_S - T_S^0 = x_B R (T_S^0)^2 / \Delta_V H.$$

Die relative Siedepunktserhöhung $\Delta T_S / T_S^0$ ist dem Molenbruch x_B des gelösten Stoffs proportional. T_S^0 ist die Siedetemperatur des Lösungsmittels, T_S die der Lösung, R die molare Gaskonstante und $\Delta_V H$ die molare Verdampfungsenthalpie des Lösungsmittels.
Da die Dampfdruckkurve der Lösung diejenige der festen Phase des Lösungsmittels erst bei einer niedrigeren Temperatur erreicht (↑ Abb. 6.3.3-7), verursacht die Dampfdruckerniedrigung in jedem Fall auch eine *Gefrierpunktserniedrigung* ΔT_E. Für diese gilt das *van't-Hoffsche Gesetz der Gefrierpunktserniedrigung*

(3) $$\Delta T_E = T_E - T_E^0 = -x_B R (T_E^0)^2 / \Delta_S H.$$

T_E ist die Erstarrungstemperatur der Lösung, T_E^0 die des reinen Lösungsmittels, und $\Delta_S H$ ist die molare Schmelzenthalpie des Lösungsmittels. Die relative Gefrierpunktserniedrigung $\Delta T_E / T_E^0$ ist dem Molenbruch des gelösten Stoffs x_B proportional und hängt ansonsten nur von der Art des Lösungsmittels, nicht aber von der Natur des Gelösten ab.
Die genannten Effekte als solche treten auch bei realen Mischphasen auf. Ihre Beschreibung erfordert aber dort eine Berücksichtigung der Aktivitäten a_B (↑ 6.1.2.4.1.) anstelle der analytischen Konzentrationen sowie der partiellen molaren Umwandlungsenthalpien.
Allgemeine praktische Bedeutung hat die Gefrierpunktserniedrigung z. B. beim Zusatz von *Gefrierschutzmitteln* zum Kühlwasser von Automotoren, beim Sprühen der Straßen mit Abfallaugen der Kaliindustrie, bei *Kältemischungen* aus Eis und Salz zur Erzeugung tiefer Temperaturen oder bei der Anwendung von *Zuschlagstoffen* bei metallurgischen Prozessen zur Herabsetzung der Schmelztemperaturen der Schlacken oder von Al_2O_3-Schmelzen.

6.3.3.3.5. *Osmotischer Druck.* Unter *Osmose* versteht man die Wanderung des Lösungsmittels einer flüssigen Mischphase durch eine nur für diese Komponente durchlässige, *semipermeable Membran*, wenn sich auf deren anderer Seite das reine Lösungsmittel befindet. Die Membran unterbindet den von allen Teilchenarten der Mischphase angestrebten Konzentrationsausgleich, indem jeweils nur die kleinsten Teilchen, d. h. die Lösungsmittelmoleküle, durch ihre Poren in Richtung des Konzentrationsgefälles diffundieren können. Nach einiger Zeit stellt sich dann ein Gleichgewichtszustand mit einem bestimmten Überdruck π der Lösung gegenüber dem reinen Lösungsmittel ein, den man als *osmotischen Druck* bezeichnet. Er hängt außer von der Temperatur vor allem von der Konzentration des Gelösten ab. Für sehr verdünnte Lösungen gilt näherungsweise $\pi = (n_B/v)\, RT = c_B RT$. Der osmotische Druck sehr verdünnter Lösungen ist also proportional der Konzentration c_B des Gelösten, aber unabhängig von dessen stofflicher Natur. Als kolligative Eigenschaft hängt er direkt mit der Dampfdruckerniedrigung zusammen. Näherungsweise ergibt sich $\pi = \Delta p RT / (p_A^0 V_A)$, der osmotische Druck ist der relativen Dampfdruckerniedrigung $\Delta p / p_A^0$ direkt proportional und kann daher aus Dampfdruckmessungen bestimmt werden (p_A^0 Dampfdruck, V_A molares Volumen des reinen Lösungsmittels bei der gleichen Temperatur T). Dieses Verfahren ist besonders für niedermolekulare Stoffe wegen der Schwierigkeit, geeignete streng semipermeable Membranen für eine direkte Messung zu finden, von Bedeutung.
Der osmotische Druck ist bereits bei relativ verdünnten Lösungen recht beträchtlich; er beträgt schon bei einer 0,1 molaren Lösung, z. B. von 34,2 g Zucker in 1 l Wasser, bei 20 °C rund $2{,}4 \cdot 10^5$ Pa (2,4 bar). Mit steigender Konzentration nimmt er meist bedeutend stärker zu als es der van't-Hoffschen Beziehung als Grenzgesetz entspricht. Direkte Bestimmungen werden vor allem zur Ermittlung der molaren Massen makromolekularer Natur- und Kunststoffe durchgeführt (↑ 5.4.3.4.1.).
Eine bedeutsame Rolle spielen osmotische Erscheinungen in biologischen Systemen. Die Wände pflanzlicher und tierischer Zellen sind für Wasser im allgemeinen durchlässig, für die meisten gelösten anorganischen und organischen Stoffe dagegen nicht.

6.3.4. Elektrolytlösungen

Als *Elektrolyte* bezeichnet man alle chemischen Stoffe, die im festen, flüssigen (geschmolzenen) oder gelösten Zustand *elektrolytisches Leitvermögen* (↑ 6.3.5.5.) aufweisen. Von den *elektronischen Leitern*, den Metallen und Halbleitern, unterscheiden sie sich dadurch, daß der Stromfluß, die *Migration*, durch positiv *und* negativ geladene atomare oder molekulare Masseteilchen, die *Ionen*, erfolgt, die zur Wahrung der Elektroneutralität in jeder elektrolythaltigen Phase in äquivalenter Menge vorhanden sein müssen.

Nach der Herkunft der Ionen in einer elektrolytischen Flüssigkeit unterscheidet man zwischen echten und potentiellen Elektrolyten. *Echte Elektrolyte* sind schon im festen Aggregatzustand aus Ionen aufgebaut, bilden also Ionengitter (↑ 6.4.3.). Bei höheren Temperaturen oder durch Einwirkung eines Lösungsmittels werden die Bindungskräfte zwischen diesen Gitterbausteinen überwunden, wodurch freie und bewegliche Ionen entstehen. Zu ihnen gehören fast alle Salze und viele Metalloxide und -hydroxide. *Potentielle Elektrolyte* leiten den Strom *nur* im gelösten Zustand, da sie erst durch eine *chemische Reaktion mit dem Lösungsmittel* Ionen bilden, z. B. $HCl + H_2O \rightarrow H_3O^+ + Cl^-$. Als Lösungsmittel können dabei auch arteigene Moleküle fungieren, z. B. $H_2O + H_2O \rightarrow H_3O^+ + OH^-$. Bei diesen Stoffen handelt es sich um kovalente Verbindungen mit starkem partiellem Ionencharakter ihrer Moleküle. Hierzu gehören alle Säuren, die meisten organischen Basen, Wasser und Ammoniak. Die Mehrzahl der echten Elektrolyte sowie die anorganischen Säuren sind *starke Elektrolyte*. In Lösungen liegen sie auch noch bei höheren Konzentrationen weitgehend, häufig praktisch vollständig in Ionen dissoziiert vor. Bei allen anderen, den *mittelstarken* und *schwachen Elektrolyten*, liegt dagegen das elektrolytische Dissoziationsgleichgewicht immer mehr oder weniger weit auf der Seite der undissoziierten Verbindungen, so daß solche Lösungen stets weniger Ionen enthalten als gleichkonzentrierte Lösungen starker Elektrolyte.

Bezüglich ihrer kolligativen Eigenschaften (↑ 6.3.3.3.) gehorchen Elektrolytlösungen den Grenzgesetzen für ideal verdünnte Lösungen, wenn die infolge der Dissoziation erhöhte Teilchenzahl berücksichtigt wird. Ursache für Abweichungen schon bei relativ niedrigen Konzentrationen sind die *interionischen Wechselwirkungen* zwischen den geladenen Teilchen, die von wesentlich größerer Stärke und Reichweite sind als bei den übrigen zwischenmolekularen Wechselwirkungen.

6.3.4.1. Bildung und Struktur von Elektrolytlösungen. Die Bildung einer Elektrolytlösung aus den reinen Komponenten hängt wesentlich von der Natur des Elektrolyten und bestimmten Eigenschaften der entstehenden Ionen, z. B. Größe und Ladung, sowie von Eigenschaften des Lösungsmittels, vor allem seiner Dielektrizitätskonstanten ε, der Größe und dem Dipolmoment μ seiner Moleküle und damit zusammenhängenden Strukturbesonderheiten ab.

6.3.4.1.1. *Auflösung echter und potentieller Elektrolyte.* Zur Überführung der Ionenbausteine eines Festkörpers in die Lösung müssen die Gitterkräfte überwunden werden. Der dazu erforderliche Energiebetrag wird meist von einem exothermen Vorgang geliefert, der in einer Anlagerung von Lösungsmittelmolekülen an die Ionen besteht und als *Solvatation*, im Fall des Lösungsmittels Wasser als *Hydratation*, bezeichnet wird.

Ursache der Solvatation sind in der Hauptsache die elektrostatischen Wechselwirkungen zwischen den Dipolmolekülen des Lösungsmittels und den Ionen. Sie sind, insbesondere bei Wasser, häufig so stark, daß sowohl die Wechselwirkungen der Lösungsmittelmoleküle untereinander als auch die Coulombschen Anziehungskräfte zwischen den Ionen weitgehend überwunden und diese dadurch als Einzelteilchen in der Lösung stabilisiert werden. Eine Flüssigkeit hat daher um so bessere Lösungsmitteleigenschaften für Elektrolyte, je größer ihre Dielektrizitätskonstante und das Dipolmoment ihrer Moleküle im Verhältnis zur Molekülgröße ist. Dies ist der Grund für die Sonderstellung des Wassers hinsichtlich seines Lösungsvermögens für Elektrolyte; die große relative Dielektrizitätskonstante $\varepsilon = 78{,}54$ bei 25 °C bewirkt eine starke Abschwächung der Ionenanziehung, und das für die geringe Molekülgröße bedeutende Dipolmoment von $\mu = 1{,}84\ D$ führt zu einer starken Ion-Dipol-Wechselwirkung und damit zu einer *Hydratationsenthalpie* in der Größenordnung der Gitterenergien. Ihre Differenz tritt nach außen als *Lösungsenthalpie* in Erscheinung und kann positiv oder negativ sein (↑ 6.3.3.2.1.).

Zur Überwindung der Wechselwirkungskräfte zwischen den Molekülen potentieller Elektrolyte sind im allgemeinen wesentlich geringere Solvatationsenthalpien erforderlich. Diese Stoffe lösen sich daher häufig auch in Lösungsmitteln mit kleinen Dipolmomenten ihrer Moleküle. In den Lösungen liegen dann aber die Elektrolyte weitgehend in undissoziierter Form vor, da die Wechselwirkungskräfte zwischen den Lösungsmittel- und den Elektrolytmolekülen nicht für eine Ladungstrennung in den gelösten Molekülen unter Bildung von Ionen ausreichen. Zur Dissoziation der Elektrolytmoleküle bedarf es vielmehr eines Lösungsmittels mit geringerer Größe und stärker pola-

rem Charakter seiner Teilchen und mit einer hinreichend großen Dielektrizitätskonstanten.

6.3.4.1.2. *Hydratation der Ionen und Struktur wäßriger Elektrolytlösungen.* Die im wesentlichen auf Ion-Dipol-Wechselwirkungen beruhende Solvatation führt zu einer *geordneten Umhüllung* der Ionen mit Lösungsmittelmolekülen. Diese ist im allgemeinen um so ausgeprägter, je kleiner der Radius und je größer die Ladung eines Ions sowie je größer das Dipolmoment der Lösungsmittelmoleküle ist. Kationen sind deshalb meist stärker solvatisiert als Anionen.

Wasserstruktur, ohne daß es allerdings wie bei kleinen oder mehrfach geladenen Teilchen zur Ausbildung einer ausgeprägten primären Hydrathülle kommt. Zur Abgrenzung von der *überwiegend geordneten Hydratation* bei letzteren spricht man im Fall der Ion-Wasser-Dipolwechselwirkungen bei einwertigen großen Ionen von einer *überwiegend ungeordneten Hydratation.*

Die Ausbildung einer primären Hydrathülle, in der die Wassermoleküle starrer gegeneinander orientiert sind und im Mittel auch wesentlich länger in den betreffenden Winkel- und Ab-

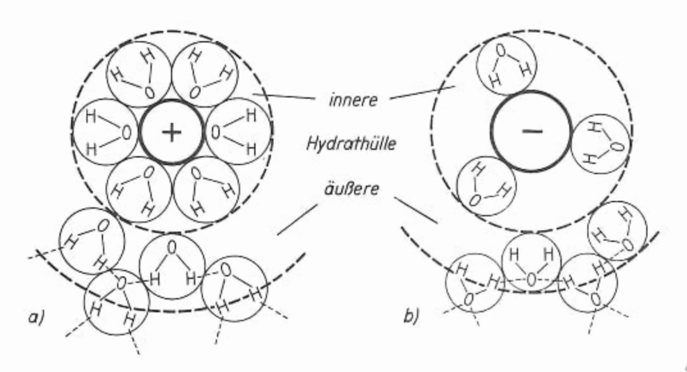


Abb. 6.3.4-1 a) hydratisiertes Kation, b) Anion, beide mit innerer und äußerer Hydrathülle (schematisch)

Bei Wasser als Lösungsmittel lagern sich an die Ionen eine bestimmte Anzahl von Wassermolekülen unter Bildung der *inneren Hydrathülle* an (Abb. 6.3.4-1). Sie sind relativ fest an das Ion gebunden und stehen mit den angrenzenden Molekülen nur noch schwächer in Wechselwirkung. Bei einer Wanderung des Ions, z. B. einer Diffusion durch Membranen oder in einem elektrischen Feld, werden die Wassermoleküle „mitgeschleppt".

Trotz der Abschirmung des elektrischen Ionenfelds durch die Dipolschicht der primären Hülle werden auch noch die angrenzenden Bereiche beeinflußt. Es kann dann noch eine *äußere Hydrathülle* entstehen, die jedoch stark dynamischen Charakter hat, indem ihre Moleküle über H-Brücken-Wechselwirkungen in ständigem Austausch mit den ungestörten Strukturelementen des Wassers stehen. Andererseits ist durch den Einfluß des elektrischen Felds der Ionen diese Wechselwirkung geringer als in der reinen Flüssigkeit, so daß diese Moleküle freier beweglich und im Mittel weniger stark zueinander orientiert sind. Im Bereich der äußeren Hydrathüllen erfolgt daher eine *Strukturbrechung* der normalen Wasserstruktur durch die gelösten Teilchen. Dieser Effekt ist besonders stark bei den großen, einfach positiv oder negativ geladenen Ionen, wie Rb^+, Cs^+, J^-, NO_3^- oder ClO_4^-. Bei diesen ergeben sich relativ weitreichende Abbruchstörungen der standslagen verbleiben als im reinen Wasser, führt zu stärker geordneten Bereichen. Man bezeichnet diese Wirkung gelöster Teilchen als *Strukturverstärkung*; sie unterscheidet sich aber von der Verstärkung der Nahordnungsstruktur in reinem Wasser bei Temperaturerniedrigung. Eine Strukturverstärkung tritt z. B. auch bei Anwesenheit von organischen Ionen mit größeren Kohlenwasserstoffgruppen oder unpolaren Molekülen in der Lösung auf. Die Wechselwirkungskräfte zwischen solchen Gruppen bzw. Teilchen und den Wassermolekülen sind sehr gering. Daher versuchen die an sie angrenzenden Wassermoleküle untereinander durch andersartige Orientierungen in stärkere Wechselwirkungen zu treten. Es kommt zur Ausbildung von Zonen verstärkter

Tab. 6.3.4-2 Molare Hydratationsenthalpien $\Delta_H H$ in kJ mol^{-1} verschiedener Ionen bei 25 °C

Ion	$\Delta_H H$	Ion	$\Delta_H H$
H^+	−1130	Ca^{2+}	−1670
Li^+	−552	OH^-	−318
Na^+	−443	F^-	−410
K^+	−402	Cl^-	−330
NH_4^+	−397	Br^-	−297
Mg^{2+}	−2000	I^-	−259

Strukturierung mit einer erhöhten Anzahl von Wasserstoffbrücken zwischen den Wassermolekülen. Diese Erscheinung wird als *Hydratation zweiter Art* bezeichnet. Sie ist auch an den Phasengrenzen von Wasser anzunehmen.

Die in den *Hydratationsenthalpien* (Tab. 6.3.4-2) zum Ausdruck kommende gesamte Wechselwirkung zwischen Ionen und Wassermolekülen umfaßt die innere und die äußere Hydratation sowie die Strukturänderungen in den angrenzenden Lösungsmittelbereichen. Deswegen sind bis heute auch die *Hydratationszahlen*, d. h. die Anzahl der in der inneren Hydrathülle enthaltenen Wassermoleküle, nicht genau bestimmbar. In Übereinstimmung mit theoretischen Abschätzungen folgen aus den Ergebnissen von experimentellen Untersuchungen mit verschiedenen Methoden, z. B. der Röntgenbeugung oder der magnetischen Kernresonanz, Hydratationszahlen von etwa 4 bis 5 für einwertige kleinere Kationen, z. B. Li^+ und Na^+, und von etwa 6 bis 8 für mehrwertige Kationen, z. B. Mg^{2+} und Al^{3+}, sowie 1 bis 3 für Cl^- und Br^-. Besonders starke Hydratationen zeigen die Ionen der Übergangsmetalle, da die einsamen Elektronenpaare des Wasser-Sauerstoffs mit deren unbesetzten Orbitalen in Wechselwirkung treten.

Spezielle Verhältnisse liegen bei den *H^+- und OH^--Ionen* in wäßrigen Lösungen vor. Freie Protonen existieren in diesen Medien nicht. Vielmehr ist jedes Proton zunächst an ein Wassermolekül durch eine Art koordinativer Bindung angelagert: $H^+ + H_2O \rightarrow H_3O^+$. Die Existenz solcher *Hydroniumionen* ist experimentell gesichert. Nach quantenmechanischen Berechnungen stellt dieses Teilchen eine flache, symmetrische Pyramide mit Valenzwinkeln zwischen 105 und 110° dar, in der die positive Überschußladung im zeitlichen Mittel gleichförmig über alle drei Protonen verteilt ist.

Die *Hydratation des Hydroniumions* führt zu einem pyramidenförmigen $H_9O_4^+$-Komplex, in dem jedes H_3O^+-Ion über H-Brücken primär mit drei Wassermolekülen verbunden ist. An den Komplex (Abb. 6.3.4-3) können weitere Moleküle unter Bildung einer äußeren Hydrathülle angelagert sein. Bei Temperatursteigerung wird diese nach und nach abgebaut, wogegen der $H_9O_4^+$-Komplex zwischen 0 und 100 °C weitgehend unbeeinflußt bleibt. Eine ähnliche innere Hydrathülle wird für das OH^--Ion angenommen. Durch Wechselwirkung mit ebenfalls drei Wassermolekülen bildet sich der Komplex $H_7O_4^-$.

Innerhalb beider Komplexe sind die positive oder negative Überschußladung längs aller H-Brücken praktisch frei beweglich. Dies erklärt die außergewöhnlich hohen Leitfähigkeiten wäßriger Säure- oder Basenlösungen (↑ 6.3.5.5.) und die sehr schnell verlaufenden Neutralisationsreaktionen: Sobald sich die $H_9O_4^+$- und die $H_7O_4^-$-Komplexe hinreichend genähert haben, erfolgt die Neutralisation zwischen den H^+- und OH^--Ionen nur noch durch spontane Weiterleitung der positiven Überschußladung aus dem $H_9O_4^+$-Komplex längs der Wasserstoffbrücken.

6.3.4.2. Theorie der interionischen Wechselwirkung. Es ist heute noch nicht möglich, die *Wechselwirkungen zwischen Ionen und Wassermolekülen* in allen Details quantitativ zu beschreiben. Für die *Wechselwirkungen gelöster Ionen untereinander* durch das weitreichende Coulomb-Potential $\varphi(r)_{\mathrm{Coul}} = -z_+z_-e_0^2/(4\pi\varepsilon_0\varepsilon r)$, wobei z_+, z_- Ionenwertigkeiten, e_0 Elementarladung, ε_0 Influenzkonstante, ε relative Dielektrizitätskonstante des Lösungsmittels, r Abstand der Ladungsschwerpunkte der beiden Ionen bedeuten, existieren dagegen, zumindest für verdünnte Lösungen, theoretisch und experimentell gut gesicherte Vorstellungen.

Ausgangspunkte der zuerst von DEBYE und HÜCKEL 1923 ausgearbeiteten Theorie sind eine Reihe vereinfachender Annahmen über den Zustand der Ionen in der Lösung: Der Elektrolyt ist vollständig dissoziiert, zwischen den Ionen wirken nur die elektrostatischen Wechselwirkungsenergien entsprechend dem Coulombpotential und diese sind klein gegenüber der Energie der Wärmebewegung, die Ionen stellen punktförmige Ladungen mit kugelsymmetrischen elektrischen Feldern dar und die Dielektrizitätskonstante der Lösung ist gleich der des reinen Lösungsmittels. Obwohl alle Ionen der regellosen Wärmebewegung unterliegen, werden sich wegen der relativ starken und weitreichenden Anziehungskräfte zwischen entgegengesetzt

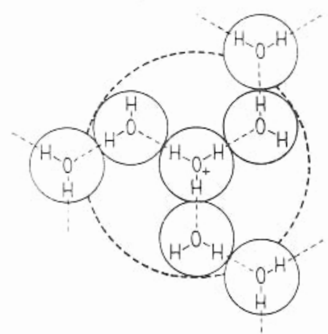


Abb. 6.3.4-3 Hydratationsmodell des Hydroniumions in wäßrigen Lösungen (schematisch)

geladenen Teilchen im zeitlichen Mittel in der Umgebung eines positiven Ions mehr negative als positive Ionen befinden und umgekehrt. Um jedes Ion bildet sich daher eine entgegen-

gesetzt geladene *Ionenwolke*, die aber nicht als längere Zeit stabiles Gebilde wie etwa eine innere Hydrathülle, sondern lediglich als statistisches Zustandsmittel einer sehr großen Zahl von Augenblicksverteilungen aufzufassen ist .Jedes Ion der Ionenwolke steht mit einem Bruchteil seiner Ladung mit dem Zentralion in Wechselwirkung.

Unter Anwendung des Boltzmannschen Energieverteilungssatzes lassen sich aus diesen Vorstellungen quantitative Beziehungen für die *interionischen Wechselwirkungsenergien* und damit die gegenseitigen Beeinflussungen der Ionen in der Lösung ableiten. Diese äußern sich hinsichtlich aller von der Konzentration abhängigen makroskopischen Eigenschaften der Lösung so, als ob jede Ionenart nur mit einem Bruchteil a ihrer analytischen Konzentration c zu den betreffenden Eigenschaften beitragen würde. Der Zusammenhang zwischen der *Aktivität* a und c ist darstellbar durch $a = fc$ (↑ 6.1.2.4.1.); der *Aktivitätskoeffizient* f der Ionen ist normalerweise kleiner als 1.

Quantitativ ergibt sich für den Aktivitätskoeffizienten f_i der Ionenart i aus der Debye-Hückel-Theorie $\ln f_i = -z_i^2 e_0^2/(8\pi\varepsilon_0\varepsilon k_B T r_w)$, wobei k_B die Boltzmannkonstante und T die Temperatur ist. r_w ist ein Faktor, der die Dimension einer Länge hat und als *Radius der Ionenwolke* oder *Debyescher Radius* bezeichnet wird. Anschaulich gibt er den Abstand vom Zentralion an, in dem ein Ion entgegengesetzter Ladung die gleiche Wirkung ausüben würde wie die Ionenwolke. Er ist um so größer, d. h. die Anreicherung entgegengesetzt geladener Ionen um das Zentralion und damit die interionische Wechselwirkung um so geringer, je niedriger die Konzentration und je höher die Temperatur sowie die Dielektrizitätskonstante des Lösungsmittels sind. Aus der Theorie folgt $r_w = [\varepsilon_0\varepsilon k_B T/(2e_0^2 I_c)]^{1/2}$. $I_c = 0{,}5\,(\sum c_i z_i^2)$ ist die *Ionenstärke*, ein zur Darstellung der interionischen Wechselwirkung in Elektrolytlösungen sehr zweckmäßiges Konzentrationsmaß. Mit ihm wird die Erfahrung berücksichtigt, daß in genügend verdünnten Elektrolytlösungen für die interionischen Wechselwirkungen nicht mehr die individuelle Art der Ionen, sondern lediglich ihre Konzentration und ihre Ladung maßgebend sind.

Für wäßrige Lösungen bei 25 °C ergibt sich aus den angeführten Beziehungen $\lg f_i = -0{,}5093 z_i^2 \sqrt{I_c}$. Da individuelle Ionenaktivitätskoeffizienten nicht zugänglich sind, definiert man mit $f_\pm = (f_+ f_-)^{1/2}$ einen *mittleren Aktivitätskoeffizienten*, der z. B. aus Dampfdruck- oder Zellspannungsmessungen ermittelt werden kann. Wie Abb. 6.3.4-4 zeigt und aus den eingangs erwähnten Näherungsannahmen verständlich wird, hat die aus der Debye-Hückel-Theorie folgende Gleichung den Charakter eines *Grenzgesetzes für sehr verdünnte Lösungen* starker Elektrolyte; bei 1–1-wertigen Elektrolyten gilt es nur bis zu einer Konzentration von 10^{-2} mol l^{-1}, bei mehrwertigen Ionen wegen des großen Einflusses der Ladungszahl nur bis etwa 10^{-3} mol l^{-1}.

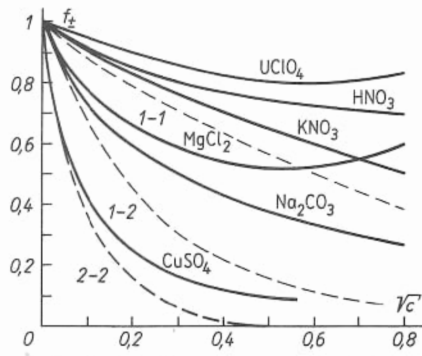


Abb. 6.3.4-4 Mittlere Ionenaktivitätskoeffizienten f_+ verschiedener starker Elektrolyte in wäßrigen Lösungen bei 25 °C; theoretischer Verlauf entsprechend dem Debye-Hückel-Grenzgesetz für 1-1-, 1-2- und 2-2-wertige Elektrolyte *gestrichelt*

Versuche zur Erweiterung des Gültigkeitsbereichs und damit zur Beschreibung der Eigenschaften konzentrierterer Elektrolytlösungen haben bisher nur zur Einführung weiterer halbempirischer Glieder geführt. Im Unterschied zu den der skizzierten Theorie zugrunde liegenden Annahmen treten in konzentrierten Lösungen starker Elektrolyte assoziierte Ionenpaare auf, werden der Hydratationszustand der Ionen und die Wasserstruktur stark beeinflußt und sind dadurch die individuellen Ioneneigenschaften von Bedeutung. Außerdem wird die elektrostatische Wechselwirkung so stark, daß sie die thermische Molekularbewegung behindert, und schließlich ergeben sich dann merkliche Veränderungen der Dielektrizitätskonstanten des Lösungsmittels in der Umgebung der Ionen.

6.3.5. Transporterscheinungen in Flüssigkeiten

Im Unterschied zu den Gasen, in denen ein *Transport von Impulsen* oder *von Energie* praktisch ausschließlich durch die Translationsbewegungen und Zusammenstöße der Teilchen bewirkt wird (↑ 6.2.5.), verläuft dieser in Flüssigkeiten im wesentlichen über *Deformationen der Nahordnungsstruktur*. Von energiereicheren Teilchen wird eine zusätzliche Kraftwirkung auf ihre Nachbarn ausgeübt und damit

der Impuls oder die Überschußenergie weitergegeben. Für einen *Massetransport* in Flüssigkeiten sind *Platzwechselvorgänge* die Ursache.

6.3.5.1. Viskosität. Auch bei Flüssigkeiten ist die *dynamische Viskosität* η (↑ 6.2.5.1.) ein Maß für die Kraft, die einer gegenseitigen Verschiebung zweier aneinandergrenzender Schichten entgegenwirkt. Ursache der dabei auftretenden Reibung sind die Wechselwirkungskräfte zwischen den Teilchen in der Grenzschicht.

Die experimentell bestimmbaren *Viskositätskoeffizienten reiner Flüssigkeiten* variieren infolge der unterschiedlichen Teilchengrößen und Wechselwirkungskräfte in weiten Grenzen (Tab. 6.3.5-1), unter Normalbedingungen von etwa 10^{-4} bis 1 $kgm^{-1}s^{-1}$ (10^{-4} bis 1 Pa · s). Da mit zunehmender Wärmebewegung die zwischenmolekularen Wechselwirkungskräfte immer weniger wirksam werden, nimmt die Viskosität von Flüssigkeiten im Unterschied zu den Gasen mit steigender Temperatur stets ab, und zwar im allgemeinen um so mehr, je höher die Absolutwerte sind und je niedriger die Temperatur ist. Die Temperaturabhängigkeit der Viskosität läßt sich empirisch durch Beziehungen der Art $\lg \eta = (A/T) + B$, mit A und B als Stoffkonstanten oder auf molekulartheoretischer Grundlage durch $1/\eta = \varphi = \varphi_0 \exp(-E_A^{vis}/RT)$ darstellen. φ wird als *Fluidität* bezeichnet und stellt gewissermaßen ein Maß für die „Leichtflüssigkeit" dar. E_A^{vis} ist die *Aktivierungsenergie des viskosen Fließens*, R die molare Gaskonstante. Der Exponentialausdruck ist demnach ein Boltzmannfaktor, der den Bruchteil der Teilchen angibt, die eine hinreichende Energie für die Überwindung des Reibungswiderstands besitzen. φ_0 ist eine stoffeigene Konstante. Die η-Abnahme mit steigender Temperatur hat große praktische Bedeutung, z. B. für die Schmierungstechnik. Druckerhöhung führt fast immer zu einer Zunahme der Viskosität, wenn nicht, wie im Wasser, spezielle Strukturverhältnisse vorliegen, die in bestimmten Druckbereichen zunächst zu einer größeren Beweglichkeit der Teilchen führen.

Tab. 6.3.5-1 Dynamische Viskositäten η in $kgm^{-1}s^{-1}$ und Wärmeleitfähigkeiten λ in $Wm^{-1}K^{-1}$ verschiedener Flüssigkeiten bei Normalbedingungen ($p_0 = 1{,}01325 \cdot 10^5$ Pa; $T_0 = 273{,}15$ K)

Flüssigkeit	$\eta \cdot 10^3$	λ
H_2O	1,787	0,551
C_2H_5OH	1,773	0,190
$C_2H_5OC_2H_5$	0,284	0,105
CH_3COCH_3	0,399	0,172
CCl_4	1,329	0,109
$CHCl_3$	0,700	0,125
C_6H_6	0,912	0,155
$C_6H_5NO_2$	3,10	0,157
$CH_2OH—CHOH—CH_2OH$	12,11	0,283
n-C_8H_{18}	0,706	0,143

Die *Viskosität flüssiger Mischphasen* läßt sich nicht allgemeingültig beschreiben; Lösungen fester Stoffe haben oft eine höhere Viskosität als das Lösungsmittel bei der gleichen Temperatur. Besonders starke Veränderungen bewirken Elektrolyte. Aus der Theorie der interionischen Wechselwirkung (↑ 6.3.4.2.) ergibt sich als *Grenzgesetz* für sehr verdünnte Lösungen $\eta_c = \eta_0 - A' \sqrt{c}$. Dabei ist η_c die Viskosität bei der Konzentration c, η_0 die des reinen Lösungsmittels und A' eine kompliziert zusammengesetzte Konstante. Qualitativ läßt sich der Einfluß der interionischen Wechselwirkungskräfte auf die Viskosität dadurch verstehen, daß beim Aneinandervorbeigleiten von Flüssigkeitsschichten zwischen diesen zusätzliche Wechselwirkungen durch die gegenseitige Deformation der Ionenwolken auftreten, die den Reibungswiderstand erhöhen.

6.3.5.2. Wärmeleitung. Im Unterschied zu den Gasen (↑ 6.2.5.2.) wird der *Energietransport* bei der Wärmeleitung in Flüssigkeiten hauptsächlich durch die zwischenmolekularen Wechselwirkungskräfte vermittelt, die den durch einen Temperaturgradienten verursachten Energieunterschied zwischen benachbarten Teilchen zum Ausgleich bringen. Unter Normalbedingungen haben die meisten Flüssigkeiten Wärmeleitfähigkeiten λ um 0,1 $Wm^{-1}K^{-1}$; Ausnahmen bilden das Wasser und andere Flüssigkeiten mit Strukturbesonderheiten (↑ Tab. 6.3.5-1). Mit steigender Temperatur nimmt λ etwa linear ab; beim Wasser erreicht sie jedoch erst bei 130 °C ein Maximum.

6.3.5.3. Diffusion. Aufgrund der größeren Wechselwirkungen zwischen den Teilchen verläuft der von selbst vor sich gehende *Massetransport* in Flüssigkeiten beträchtlich langsamer als in Gasen (↑ 6.2.5.3.). Den in *reinen Flüssigkeiten* durch die thermische Molekularbewegung verursachten ungerichteten *Platzwechsel* der Atome oder Moleküle bezeichnet man als *Selbstdiffusion*. Praktisch untersuchen läßt sich diese mit radioaktiven Isotopen der betreffenden Teilchen. Jedes Flüssigkeitsteilchen führt unter Normalbedingungen etwa 10^{13} anharmonische Schwingungen je Sekunde aus; nach etwa 10^4 bis 10^5 solcher Schwingungen erfolgt im Durchschnitt ein Platzwechsel, und zwar um eine Strecke, die ungefähr einem Teilchendurchmesser entspricht. Damit „springt" jedes Teilchen in der Sekunde etwa 10^8mal in eine benachbarte Lage. Durch die fortwährenden Richtungsänderungen ist aber trotz der hohen Platzwechselzahl die Verschiebung in eine *bestimmte* Richtung nur gering.

Dies gilt im gleichen Maße auch für die gelösten Teilchen in einer *flüssigen Mischphase*. Daher erfordert auch beim Vorliegen eines Konzentrationsgefälles der Konzentrationsausgleich innerhalb der gesamten Mischphase eine erhebliche Zeit, wenn dieser nicht durch Konvektion der Flüssigkeit unterstützt wird. Unter Normalbedingungen liegen die Werte der Diffusionskoeffizienten D gelöster Stoffe in flüssigen Phasen etwa zwischen 0,5 und $4 \cdot 10^{-5}$ cm^{2}s^{-1}. Sie hängen ab von der Größe, der Solvatation und anderen Eigenschaften der diffundierenden Teilchen, von verschiedenen Eigenschaften des Lösungsmittels, in erster Linie seiner Viskosität η, sowie von der Temperatur. Bei annähernd kugelförmigen Teilchen vom Radius r ist der Reibungswiderstand, den sie bei ihrer Bewegung durch das betreffende Medium überwinden müssen, näherungsweise nach dem Stokesschen Gesetz durch $6\pi\eta r$ gegeben. Für den Diffusionskoeffizienten gilt dann $D = k_B T/(6\pi\eta r)$, wobei k_B die Boltzmannkonstante ist. Mit der Temperatur steigt D stets merklich an, zumal bei Flüssigkeiten dabei die Viskosität noch abnimmt. Eine genauere Darstellung der *Temperaturabhängigkeit des Diffusionskoeffizienten* gelöster Teilchen ist durch $D = D_0 \exp(-E_A^{\text{diff}}/RT)$ möglich. Hierin bedeuten E_A^{diff} eine *Aktivierungsenergie für die Platzwechselvorgänge*, d. h. für das „ruckweise Hindurchzwängen" durch die Masse der Lösungsmittelteilchen, R die molare Gaskonstante und D_0 eine stoffeigene Konstante.

Mit steigender Konzentration erniedrigen sich die D-Werte oftmals beträchtlich, was durch die Zunahme der Wechselwirkung der gelösten Teilchen *untereinander* verständlich wird. Außerdem spielen dabei auch Veränderungen der Solvatation und der Viskosität der Mischphase eine Rolle. In Elektrolytlösungen diffundieren die Kationen und die Anionen mit gleicher Geschwindigkeit in Richtung eines Konzentrationsgefälles, wodurch die Elektroneutralität auch in kleinen Bereichen der Mischung immer gewahrt bleibt.

6.3.5.4. Schallausbreitung. Schallwellen breiten sich in Flüssigkeiten wie in Gasen (↑ 6.2.5.4.) in Form von Longitudinalwellen aus, wobei die Druckänderungen im Ultraschallbereich adiabatisch erfolgen. Für die *Schallgeschwindigkeit* gilt $w_{Sch} = \sqrt{1/(\varrho\beta_{ad})} = \sqrt{\varkappa/(\varrho\beta_{is})}$; ϱ ist die Dichte, β_{ad} die adiabatische, β_{is} die isotherme Kompressibilität und $\varkappa = \beta_{is}/\beta_{ad} = C_p/C_v$ der Poissonsche Koeffizient. Schallgeschwindigkeitsmessungen liefern daher die Möglichkeit zur Bestimmung der experimentell nicht direkt zugänglichen Molwärme C_v. Darüber hinaus gestatten sie Aussagen über die Wechselwirkungskräfte zwischen den Flüssigkeitsteilchen, über die der *Energietransport* bei der Schallausbreitung erfolgt.

Unter Normalbedingungen liegen die Schallgeschwindigkeiten in den meisten *reinen Flüssigkeiten* zwischen 800 und 2000 ms^{-1}, in Wasser ist $w_{Sch} = 1483{,}1$ ms^{-1} bei 20 °C. Mit steigender Temperatur nimmt w_{Sch} im allgemeinen fast linear um etwa 0,4%K^{-1} ab. Wasser zeigt auch hier seine Ausnahmestellung: die Schallgeschwindigkeit steigt von 0 °C zunächst parabolisch an, erreicht bei 73,4 °C mit 1555,5 ms^{-1} einen Maximalwert, um danach wie bei allen „normalen" Flüssigkeiten abzufallen. Die Erklärung ist wiederum in den Strukturbesonderheiten zu sehen; mit steigender Temperatur wird über den Abbau der sperrigen Cluster eine höhere Packungsdichte der Wassermoleküle mit stärkerer gegenseitiger Kopplung für den Energietransport erreicht. Eine gleiche Wirkung haben Elektrolyte im Wasser; die dabei mit wachsender Konzentration auftretende Strukturbrechung durch die Ionen führt zu einem Ansteigen der Schallgeschwindigkeit.

Die *Schallabsorption* ist wesentlich geringer als in Gasen und die Reichweite von Schallsignalen dadurch beträchtlich größer. Während z. B. bei einer Frequenz von 100 kHz die Schallintensität in Luft bereits nach etwa 220 m auf die Hälfte des Ausgangswerts abgesunken ist, beträgt die entsprechende Entfernung im Wasser etwa 400 km. Hierauf beruht auch die verbreitete Anwendung von Ultraschall in der Schiffahrt für Echolotungen von Seetiefen und Ortung von Fischschwärmen.

6.3.5.5. Elektrolytische Leitfähigkeit. Beim Anlegen einer Gleichspannung an die Elektroden einer elektrolytischen Zelle (↑ 7.3.4.) entsteht zwischen ihnen ein elektrisches Feld der Stärke $E = U/l$, wobei l der Elektrodenabstand und U die Spannung über der Elektrolytlösung ist. In diesem Feld wandern die positiven und negativen Ionen mit einer von der Feldstärke, der Art des Elektrolyten und des Lösungsmittels, der Konzentration des Elektrolyten und der Temperatur abhängigen Geschwindigkeit zu den entgegengesetzt gepolten Elektroden; das elektrische Feld bewirkt einen *gerichteten Massetransport*, die *Migration*.

6.3.5.5.1. *Kenngrößen der elektrolytischen Leitfähigkeit.* Wie bei metallischen Leitern versteht man auch bei Elektrolytlösungen unter der *elektrischen* bzw. *elektrolytischen Leitfähigkeit* den Kehrwert des elektrischen Widerstands R. Dieser hängt nach $R = \varrho l/A$ von der Länge l und dem Querschnitt A des Leiters ab. Der Proportionalitätsfaktor ϱ ist eine stoff- und temperaturabhängige, bei Elektrolytlösungen noch konzentrationsabhängige Materialkonstante und heißt *spezifischer Widerstand.* Sein Kehrwert $1/\varrho = \varkappa$ ist die *spezifische Leitfähig-*

keit; bei Elektrolytlösungen gibt sie nach $\varkappa = l/(RA)$ die Leitfähigkeit $1/R$ eines cm^3 der Lösung zwischen zwei Elektroden von je 1 cm^2 Fläche im Abstand von 1 cm an. Ihre Einheit ist damit $\Omega^{-1}cm^{-1}$.
Die spezifischen Leitfähigkeiten von Elektrolytlösungen liegen etwa zwischen 10^{-8} und $1\ \Omega^{-1}cm^{-1}$ (Abb. 6.3.5-2); für reinstes und völlig CO_2-freies Wasser beträgt sie bei 25 °C $3{,}73 \cdot 10^{-8}\ \Omega^{-1}cm^{-1}$.

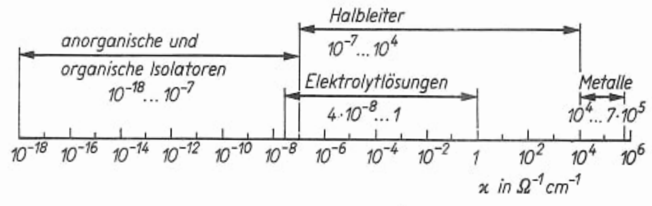


Abb. 6.3.5-2 Spezifische Leitfähigkeiten $\varkappa$ verschiedener Stoffgruppen bei Normaltemperatur

Zum Vergleich der Leitfähigkeiten von Elektrolytlösungen untereinander benutzt man die *Äquivalentleitfähigkeit* $\Lambda = \varkappa/c_{ev}$ mit der Maßeinheit $cm^2\Omega^{-1}val^{-1}$. Die Äquivalentkonzentration ist $c_{ev} = z_e c$; z_e ist die elektrochemische Wertigkeit des Elektrolyten, die sich nach $z_e = \nu_+ z_+ = \nu_- z_-$ aus der Anzahl ν_+ der Kationen mit der Ladungszahl z_+ und der Anzahl ν_- der Anionen mit der Ladungszahl z_- ergibt, in die der betreffende Elektrolyt dissoziiert.

Im elektrischen Feld bewegen sich die Ionen mit einer der Feldstärke direkt und dem Reibungswiderstand indirekt proportionalen, stationären Geschwindigkeit. Die von der Art des Ions abhängigen Geschwindigkeiten werden miteinander vergleichbar, wenn man sie auf die Einheitsfeldstärke $1\ Vcm^{-1}$ bezieht. Diese speziellen Wanderungsgeschwindigkeiten heißen *Ionenbeweglichkeiten* u_+ bzw. u_-; ihre Einheit ist $cm^2V^{-1}s^{-1}$.
In sehr verdünnten wäßrigen Lösungen, in denen sich die Ionen bei ihrer Migration gegenseitig nicht behindern, liegen die Beweglichkeiten der meisten Ionen bei normaler Temperatur zwischen 4 und $8 \cdot 10^{-4}\ cm^2V^{-1}s^{-1}$. Eine Ausnahme stellen die H^+- und OH^--Ionen dar, mit $u_{H^+} = 36{,}3 \cdot 10^{-4}\ cm^2V^{-1}s^{-1}$ sowie $u_{OH^-} = 20{,}5 \cdot 10^{-4}\ cm^2V^{-1}s^{-1}$ bei 25 °C. Dafür ist ein durch nur geringfügige Verschiebungen von Protonen in den H-Brücken der hydratisierten $H_9O_4^+$- bzw. $H_7O_4^-$-Komplexe (↑ 6.3.4.1.2.) vermittelter Transport positiver Ladungen die Ursache. Da andererseits die Protonenbeweglichkeit im Eisgitter noch etwa 50mal größer als in Wasser bei 0 °C ist, wird als geschwindigkeitsbestimmender Teilvorgang des Ladungstransports in der flüssigen Phase die Neuformierung der $H_9O_4^+$- bzw. $H_7O_4^-$-Komplexe durch Bildung neuer H-Brücken an der vorderen und Abbau solcher Brücken an der rückseitigen Peripherie der Komplexe in den entsprechenden Feldrichtungen angesehen.
Die gesamte Leitfähigkeit einer Elektrolytlösung setzt sich stets additiv aus den Beiträgen des Ladungstransports durch die Kationen und die Anionen zusammen. Sie ist deshalb um so größer, je höher die Konzentration der Ionen in der Lösung, je größer ihre Ladungszahlen und je höher die Beweglichkeiten beider Ionenarten sind. Für die spezifische Leitfähigkeit gilt $\varkappa = c_{ev}F\alpha(u_+ + u_-)$; bei Lösungen starker Elektrolyte ist der Dissoziationsgrad α gleich Eins. Die Faradaykonstante lautet
$F = 96487\ As\,mol^{-1}$. Die Produkte Fu_+ bzw. Fu_- bezeichnet man als *Ionenäquivalentleitfähigkeiten* Λ_+ bzw. Λ_-; ihre Einheit ist $cm^2\Omega^{-1}val^{-1}$. Damit gilt für die Äquivalentleitfähigkeit Λ einer Lösung $\Lambda = \alpha(\Lambda_+ + \Lambda_-)$.

6.3.5.5.2. *Konzentrations- und Temperaturabhängigkeit der elektrolytischen Leitfähigkeit.* Bei niedrigen Konzentrationen ist die *spezifische Leitfähigkeit* $\varkappa$ von Elektrolytlösungen näherungsweise der Konzentration proportional. Bei starken Elektrolyten wird aber mit steigender Konzentration die interionische Wechselwirkung immer stärker und bei schwachen Elektrolyten verringert sich dabei der Dissoziationsgrad. Daher durchläuft $\varkappa$ bei allen Elektrolytlösungen in Abhängigkeit von der Konzentration ein Maximum.
Die *Äquivalentleitfähigkeit* Λ nimmt dagegen stets mit abnehmender Konzentration zu und erreicht bei sehr hohen Verdünnungen einen stoffeigenen Grenzwert, den man *Äquivalentleitfähigkeit bei unendlicher Verdünnung* Λ_∞ nennt. In diesem Zustand ist die Migration jedes Ions unabhängig vom Gegenion – es gilt das *Gesetz der unabhängigen Ionenwanderung* von Kohlrausch – und allgemein $\Lambda_\infty = \Lambda_{+\infty} + \Lambda_{-\infty}$. Für die Konzentrationsabhängigkeit der Äquivalentleitfähigkeit starker Elektrolyte ($\alpha = 1$) ist bei Konzentrationen unterhalb etwa $10^{-2}\ mol\,l^{-1}$ das ebenfalls von Kohlrausch empirisch gefundene *Quadratwurzelgesetz* $\Lambda = \Lambda_\infty - k\sqrt{c}$ gültig, in dem k eine von der Art der Elektrolytlösung abhängige Konstante darstellt. Dieses Gesetz kann dazu benutzt werden, um die Λ_∞-Werte der Elektrolyte durch lineare Extrapolation der bei ver-

schiedenen Konzentrationen experimentell bestimmten Λ-Werte zu erhalten.
Zur Ermittlung der Leitfähigkeiten von Elektrolytlösungen mißt man deren elektrischen Widerstand unter Verwendung einer *Wheatstoneschen Brückenschaltung.* Diese Messungen werden mit Wechselstrom von etwa 1 kHz durchgeführt, um Stoffabscheidungen und Polarisationserscheinungen (↑ 7.3.4.2.2.) an den Elektroden zu vermeiden. Zusammen mit den ebenfalls experimentell zugänglichen Leitfähigkeitsbeiträgen einzelner Ionenarten zur Gesamtleitfähigkeit, den *Überführungszahlen,*

Tab. 6.3.5-3 Ionenäquivalentleitfähigkeiten bei unendlicher Verdünnung $\Lambda_{+\infty}$ und $\Lambda_{-\infty}$ in $cm^2\Omega^{-1}val^{-1}$ verschiedener Ionen bei 25 °C

Kation	$\Lambda_{+\infty}$	Anion	$\Lambda_{-\infty}$
H^+	349,8	OH^-	197,6
Li^+	38,7	F^-	55,0
Na^+	50,1	Cl^-	76,3
K^+	73,5	Br^-	78,4
NH_4^+	73,4	I^-	76,8
Mg^{2+}	53,1	NO_3^-	71,4
Ca^{2+}	59,5	SO_4^{2-}	79,8

sind aus den Λ_∞-Werten auch die Ionenäquivalentleitfähigkeiten bei unendlicher Verdünnung zugänglich (Tab. 6.3.5-3). In deren Werten spiegeln sich deutlich die Unterschiede im Hydratationszustand (↑ 6.3.4.1.2.) und damit in den Beweglichkeiten der einzelnen Ionen wider.
Für die *Temperaturabhängigkeit der spezifischen* und *der Äquivalentleitfähigkeit* von Elektrolytlösungen ist in erster Linie die exponentielle Viskositätsabnahme der Lösung mit steigender Temperatur maßgebend. Daher nehmen $\varkappa$ und Λ zunächst immer zu. Diesem Einfluß wirken aber eine Reihe anderer Faktoren entgegen, so daß bei wäßrigen Elektrolytlösungen die Leitfähigkeit etwa oberhalb 70 bis 90 °C wieder abfällt. Als Hauptursache hierfür ist die starke Abnahme der Dielektrizitätskonstante des Lösungsmittels mit zunehmender Temperatur anzusehen, die zu einer beträchtlichen Verstärkung der interionischen Wechselwirkungen und damit zu einer Herabsetzung der Beweglichkeiten der Ionen führt.

6.3.5.5.3. *Leitfähigkeit und interionische Wechselwirkung.* Aus der Debye-Hückel-Theorie der interionischen Wechselwirkung (↑ 6.3.4.2.) sind auch Aussagen über deren Einflüsse auf die Ionenbeweglichkeiten und damit über die Konzentrationsabhängigkeit der Äquivalentleitfähigkeiten Λ im Bereich niedriger Konzentrationen möglich.

Danach muß die Ionenwolke bei einer Wanderung des Zentralions in einem elektrischen Feld ständig neu formiert werden, weil sich deren Gegenionen in entgegengesetzte Richtung bewegen. Dies erfordert eine bestimmte Zeit, die *Relaxationszeit,* und deshalb hat die Ionenwolke bei der Migration des Zentralions eine asymmetrische Form. Als Folge dieses *Relaxationseffekts* erfährt das Zentralion durch den in Bewegungsrichtung hinter ihm liegenden Teil der Ionenwolke eine zusätzliche Bremsung. Da die Ionen in der Lösung stets mit ihren Hydrathüllen wandern, bewegen sich die entgegengesetzt geladenen Ionen außerdem nicht relativ zu einem ruhenden Medium, sondern gegen eine Art Strömung des Lösungsmittels. Die hieraus resultierende Wanderungshemmung hängt von der Viskosität η des Lösungsmittels ab und wird als *elektrophoretischer Effekt* bezeichnet.
Beide *Bremseffekte* werden um so stärker, je höher die Konzentration c und je niedriger die Temperatur T sowie die Dielektrizitätskonstante ε des Lösungsmittels ist. Sie bewirken demzufolge eine Abnahme der Äquivalentleitfähigkeit Λ mit steigender Konzentration. Die quantitative Auswertung dieser Vorstellungen ergibt für Lösungen 1–1-wertiger Elektrolyte im Bereich niedriger Konzentrationen die *Debye-Hückel-Onsager-Gleichung*

$$(4) \quad \Lambda = \Lambda_\infty - \frac{8{,}2 \cdot 10^5 \Lambda_\infty}{(\varepsilon T)^{3/2}} \sqrt{c} - \frac{82{,}48}{\eta(\varepsilon T)^{1/2}} \sqrt{c}.$$

Das zweite Glied auf der rechten Seite bringt die Auswirkungen des Relaxationseffekts und das dritte Glied des elektrophoretischen Effekts zum Ausdruck. Für ein gegebenes Lösungsmittel und einen bestimmten Elektrolyten werden bei konstanter Temperatur alle Größen dieser beiden Glieder bis auf c konstant und gleich der Konstanten k im Quadratwurzelgesetz von KOHLRAUSCH.
Bei starken Elektrolyten kann der experimentell leicht zugängliche Quotient $\Lambda/\Lambda_\infty = f_\Lambda$ als quantitatives Maß für die interionischen Wechselwirkungen betrachtet werden; f_Λ heißt *Leitfähigkeitskoeffizient* und nähert sich mit abnehmender Konzentration immer dem Wert 1. Er kann bei Kenntnis der Konstanten des Relaxations- und des elektrophoretischen Effekts auch aus diesen berechnet werden. Die gute Übereinstimmung der auf beiden Wegen erhaltenen Ergebnisse bestätigt die physikalische Realität der Modellvorstellungen, die der Theorie der interionischen Wechselwirkung und der Leitfähigkeitstheorie verdünnter Lösungen starker Elektrolyte zugrunde liegen. Für Lösungen schwacher Elektrolyte ist der Zusammenhang $\Lambda/\Lambda_\infty = \alpha f_\Lambda$ gültig. In diesem

Fall kann aus Leitfähigkeitsmessungen nur das Produkt αf_A erhalten und keine Aussage über die Stärke der interionischen Wechselwirkung allein gemacht werden.
Einer Erweiterung der Leitfähigkeitstheorie auf das Gebiet höherer Konzentrationen stehen die gleichen Probleme entgegen, wie sie schon bei der Theorie der interionischen Wechselwirkung angedeutet wurden. Bei höheren Konzentrationen nimmt die Äquivalentleitfähigkeit im allgemeinen langsamer ab, als es nach 6.3.(4) zu erwarten ist.

6.3.6. Quantenflüssigkeiten

Unter Quantenflüssigkeiten werden die flüssigen Heliumisotope He-3 und He-4 und ihre Mischungen verstanden. Aufgrund der geringen van-der-Waalsschen Wechselwirkung der Atome in diesen Flüssigkeiten im Vergleich zu ihrer Nullpunktsenergie (s. a. 6.4.) werden diese auch bis zu $T = 0$ K unter ihrem eigenen Dampfdruck nicht fest. Die Existenz der flüssigen Phase bis zu beliebig tiefen Temperaturen ist ein typischer makroskopischer Quanteneffekt, der sich auch in anderen physikalischen Eigenschaften dieser Flüssigkeiten ausdrückt.
6.3.6.1. Quantenflüssigkeit He-4. Das He-4-Atom enthält zwei Hüllenelektronen mit dem Gesamtdrehimpuls Null und in seinem Kern zwei Protonen und zwei Neutronen mit dem resultierenden Kernspin 0. Danach gehorcht es der Bose-Einstein-Statistik (↑ 2.4.4.), es können also beliebig viele Atome auch bezüglich ihrer Translationsenergie den tiefsten Energiezustand einnehmen. Das führt zu einer Reihe ungewöhnlicher Eigenschaften. Unter Normaldruck kondensiert He-4 bei 4,2 K. Bei einer Temperatur von 2,18 K, dem sogenannten λ-Punkt, geht es in eine neue flüssige Phase über, die als He II bezeichnet wird. Die auffälligste Eigenschaft von He II, die zuerst von KAPIZA entdeckt wurde, ist, daß He II seine Zähigkeit verliert, also *suprafluid* wird. Daher kann es praktisch reibungsfrei durch engste Kapillaren und Spalte fließen. Einmal in He II erzeugte Wirbel können nahezu unbegrenzt aufrechterhalten werden. Die Wärmeleitfähigkeit erreicht gegenüber normalem flüssigem He-4 10^8fach höhere Werte. Die Wärmekapazität von He II nimmt unterhalb von 2,18 K zunächst proportional T^6 ab und unterhalb von 0,6 K proportional T^3. Damit besitzt He II die gleiche Temperaturabhängigkeit der Wärmekapazität bei tiefen Temperaturen wie kristalline Festkörper.
Phänomenologisch läßt sich das Verhalten von He II durch das sogenannte *Zweiflüssigkeitsmodell* beschreiben. Danach ist He II ein Gemisch einer normalen und einer suprafluiden Komponente. Unterhalb des λ-Punkts wandelt sich mit abnehmender Temperatur immer mehr normales He-4 in suprafluides He-4 um, bis bei $T = 0$ K nur noch die suprafluide Komponente vorliegt. Die beiden Komponenten können reibungsfrei nebeneinander strömen. Die suprafluide Komponente besitzt die Entropie Null. Die normale Komponente fließt immer nach Stellen niedriger Temperatur und wandelt sich dort unter Abgabe von Wärme in die suprafluide Komponente um. Diese wiederum fließt zu Stellen höherer Temperatur, um sich dort unter Aufnahme von Wärme in normales He-4 umzuwandeln. So entsteht die große Wärmeleitfähigkeit von He II. Das Modell erklärt auch den sogenannten *Fontäneeffekt*: Bei zwei durch sehr enge Poren verbundenen Behältern unterschiedlicher Temperaturen kann nur die suprafluide Komponente in den Behälter mit etwas höherer Temperatur fließen, wo sie sich in normales He-4 umwandelt und wegen der nun endlichen Zähigkeit nicht mehr durch die engen Poren zurückfließen kann. Auf diese Weise entsteht ein Überdruck, der das Helium als Fontäne über ein Rohr aus diesem Behälter herausdrückt. In He II können sich zwei Arten von longitudinalen Wellen ausbreiten: einerseits Dichtewellen, die der normalen Schallausbreitung in Flüssigkeiten entsprechen und als erster Schall bezeichnet werden, zum anderen durch periodisches Heizen angeregte Temperaturwellen mit Frequenzen bis zu 10^2 Hz, die sich praktisch ungedämpft ausbreiten und als zweiter Schall bezeichnet werden. Er wurde aufgrund des Zweiflüssigkeitsmodells von TISZA und LANDAU vorhergesagt und von PESHKOV nachgewiesen. Er entsteht dadurch, daß periodische Temperaturänderungen in He II Konzentrationsänderungen von normaler und suprafluider Komponente hervorrufen. Da sich beide Komponenten reibungslos nebeneinander bewegen können, pflanzen sich diese Konzentrationsschwankungen und die ihnen entsprechenden Temperaturschwankungen als ungedämpfte Temperaturwelle mit Geschwindigkeiten bis zu 180 ms^{-1} fort.
Die mikroskopische Theorie zur Erklärung des Verhaltens von He II wurde vor allem durch LANDAU entwickelt.
6.3.6.2. Quantenflüssigkeit He-3. Das leichte Isotop des Heliums enthält in seinem Kern zwei Protonen, aber nur ein Neutron. Damit bleibt ein Nukleonenspin unkompensiert, und He-3-Atome haben den Spin 1/2. Folglich gehorchen sie der Fermi-Dirac-Statistik. Ähnlich wie bei den Elektronen in einem Metall kann entsprechend dem Pauliprinzip ein Energiezustand nur von jeweils zwei Teilchen mit entgegengesetzten Spins eingenommen werden. So ist zu erwarten, daß die physikalischen Eigen-

schaften der Fermiflüssigkeit He-3 prinzipiell andere sein werden als die von flüssigem He-4. Dies drückt sich schon in den thermodynamischen Eigenschaften aus. He-3 siedet unter Normaldruck bei 3,2 K. Die Wärmekapazität hängt unterhalb von 0,2 K linear von der Temperatur ab, ebenso wie der Beitrag der Elektronen zur Wärmekapazität in Metallen. Ungewöhnlich ist auch das Phasendiagramm flüssig – fest von He-3. Um den festen Zustand zu erreichen, muß flüssiges He-3 ebenso wie flüssiges He-4 unter Druck gesetzt werden. Der notwendige Druck hängt von der Temperatur ab, wird aber bei sehr tiefen Temperaturen temperaturunabhängig. Das Ungewöhnliche in der Schmelzdruckkurve von flüssigem He-3 ist ein Minimum bei $T_{min} = 0{,}319$ K und $p_{min} = 2{,}931$ MPa. Dieses Minimum ist recht tief, denn der Schmelzdruck bei $T \rightarrow 0$ K beträgt etwa 3,44 MPa. Wichtige thermodynamische Konsequenzen dieses Verhaltens folgen aus der Clausius-Clapeyronschen Gleichung, die für den Übergang fest – flüssig den Anstieg der Schmelzdruckkurve angibt zu $\mathrm{d}p/\mathrm{d}T = (S_{fl} - S_{fest})/(V_{fl} - V_{fest})$. Da entlang der Schmelzdruckkurve für die Volumina $V_{fl} < V_{fest}$ gilt, bedeutet der negative Anstieg der Schmelzdruckkurve unterhalb ihres Minimums, daß für die Entropien $S_{fl} < S_{fest}$ gelten muß. In diesem Druck-Temperaturbereich ist also die Entropie des flüssigen He-3 kleiner als die des festen He-3. Dies ist ungewöhnlich und weist darauf hin, daß die flüssige Phase des He-3 hier geordneter ist als die feste Phase. Für $T \rightarrow 0$ K wird $\mathrm{d}p/\mathrm{d}T = 0$, da aufgrund des Nernstschen Wärmetheorems (↑ 6.1.2.3.1.) $S_{fl}(0) = S_{fest}(0) = 0$. Die Abb. 6.3.6-1 zeigt das Entropiediagramm von He-3 entlang der Schmelzdruckkurve bis etwa 300 mK. In Übereinstimmung mit der obigen Analyse der Schmelzdruckkurve sind die Entropien des flüssigen und des festen He-3 bei T_{min} gleich. Bei $T_0 = 2$ mK tritt ein starker Abfall der Entropie auf, der auf einen Phasenübergang im festen He-3 hinweist. Bei dieser Temperatur ordnen sich die He-3-Kerne im Gitter antiferromagnetisch.

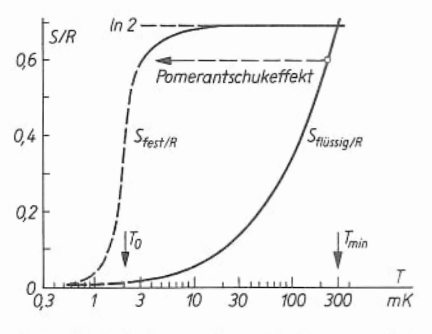


Abb. 6.3.6-1 Entropie von flüssigem und festem He-3 entlang der Schmelzdruckkurve; die gesamte Kernspinentropie ist $S/R = \ln(2I + 1) = \ln 2$, wobei R die ideale Gaskonstante ist, T_0 bedeutet die antiferromagnetische Phasenübergangstemperatur in festem He-3; die Wirkungsweise des Pomerančukeffekts (isotroper Kompression) ist angedeutet

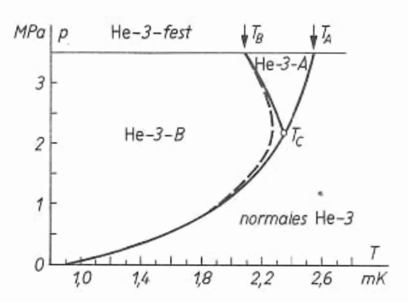


Abb. 6.3.6-2 Phasendiagramm von He-3; die neu entdeckten flüssigen Phasen He-3-A und He-3-B sind suprafluid

Bemerkenswert ist, daß zwischen 10 mK und 0,3 K die Entropie des festen He-3 nahezu konstant ist. Dieses ungewöhnliche Verhalten von He-3 ist die Grundlage der von Pomerančuk vorgeschlagenen Methode zur Erzeugung tiefer Temperaturen, die unter dem Namen *Pomerančukeffekt* bekannt ist. Dazu wird flüssiges He-3 in einer auf etwa 25 mK vorgekühlten abgeschlossenen Kammer bei thermischer Isolation, also bei konstanter Entropie (↑ Abb. 6.3.6-1) zusammengedrückt. Dabei wird flüssiges He-3 in festes He-3 umgewandelt, und das System kühlt sich auf Temperaturen ab, die unterhalb von 3 mK liegen. Diese Methode der Erzeugung sehr tiefer Temperaturen im Bereich von 2 bis 25 mK ist sehr effektiv und kann quasikontinuierlich gestaltet werden. Mit Hilfe solcher Pomerančukzellen wurde nach 1972 das flüssige He-3 untersucht und eine ganze Reihe neuer physikalischer Zusammenhänge gefunden.

6.3.6.3. Phasendiagramm und Suprafluidität von flüssigem He-3. Der heute bekannte Teil des Phasendiagramms von flüssigem He-3 ist in Abb. 6.3.6-2 dargestellt. Dieses Diagramm wurde aus Messungen der Wärmekapazität, der magnetischen Suszeptibilität, der Schallabsorption, der Kernresonanz und der Wärmeleitfähigkeit gewonnen. Wie ersichtlich, existieren neben der festen Phase drei verschiedene flüssige Phasen von He-3: die normale Fermiflüssigkeit, die sogenannte *A*-Flüssigkeit und die *B*-Flüssigkeit. Die Existenzgrenze zwischen der Fermiflüssigkeit und der *A*- und *B*-Phase beginnt bei $T_A = 2{,}6$ mK auf der Schmelzkurve

und verschiebt sich mit abnehmendem Druck nach tiefen Temperaturen. Bei $T_c = 2{,}4$ mK und $p_c = 2{,}15$ MPa existiert ein Tripelpunkt. Mit weiter abnehmendem Druck verschiebt sich der Übergang B-Flüssigkeit → normale Flüssigkeit nach tiefen Temperaturen und erreicht etwa 0,8 mK, wenn der äußere Druck auf Null reduziert wird. Der Phasenübergang von normaler Fermiflüssigkeit zur A- oder B-Phase ist zweiter Ordnung, während der Phasenübergang von A- zu B-Flüssigkeit erster Ordnung ist, da hier eine kleine Umwandlungswärme auftritt. Weiter wurde festgestellt, daß schon ein relativ kleines magnetisches Feld von 38 mT den Phasenübergang zwischen A- und B-Flüssigkeit beeinflußt, wobei der Tripelpunkt verschwindet, wie dies in Abb. 6.3.6-2 durch die gestrichelte Linie angedeutet ist. Experimentell konnte einwandfrei nachgewiesen werden, daß die A- und B-Phasen des He-3 suprafluid sind. Dies war bisher nur von flüssigem He-4 bekannt. Nach heutigen theoretischen Vorstellungen liegt der Suprafluidität der Fermiflüssigkeit He-3 ein ähnlicher Mechanismus zugrunde, wie bei der auf der Bildung von Cooperpaaren beruhenden Supraleitfähigkeit des Elektronensystems in einem Metall (↑ 8.5.3.2.). In flüssigem He-3 bilden die Atome nämlich ebenfalls Cooperpaare. Während bei den Metallen die Spins der beiden korrelierten Elektronen immer entgegengesetzt sind, können im flüssigen He-3 die paarbildenden He-3-Atome bezüglich ihres Kernspins in drei verschiedenen Konfigurationen existieren, die alle suprafluid sind. Suprafluides He-3 kann als eine Mischung einer normalen und einer oder mehrerer der suprafluiden Konfigurationen angesehen werden. Das Verhältnis der normalen zur suprafluiden Komponente in der Flüssigkeit ist eine Funktion der Temperatur. Die suprafluiden Komponenten bilden sich am Punkt T_A des Phasendiagramms und dominieren mit $T \to 0$ K. Alle suprafluiden Komponenten können teilweise nebeneinander existieren. In der A-Phase sind die Kernspins innerhalb eines Cooperpaares parallel gerichtet. Die Untersuchung vor allem der dynamischen Eigenschaften läßt noch weitere sehr interessante Ergebnisse erwarten.

6.3.6.4. He-3-He-4-Lösungen, Lösungskältemaschine. Ein bedeutsames Verhalten beobachtet man bei flüssigen Lösungen der beiden Heliumisotope. Es ergibt sich ein Phasendiagramm, wie es in Abb. 6.3.6-3 dargestellt ist. Oberhalb von 0,8 K sind flüssiges He-3 und He-4 in jedem Verhältnis mischbar. Dabei verschiebt sich, ausgehend von reinem He-4, der λ-Punkt mit wachsender He-3-Konzentration nach tiefen Temperaturen. Diese λ-Linie trennt die Existenzbereiche der normalen und suprafluiden Phase der Lösung und mündet bei 0,86 K in einen Tripelpunkt. Bei Abkühlung einer Lösung unterhalb dieser Temperatur tritt eine Entmischung der Lösung auf, und entsprechend der eingezeichneten Koexistenzkurve spaltet die Flüssigkeit in zwei Phasen auf: in eine He-3-reiche und eine He-4-reiche Phase, wobei sich die He-3-reiche Phase wegen ihrer geringeren Dichte immer oberhalb der He-4-reichen Phase befindet. Das Besondere und für die Erzeugung tiefer Temperaturen Wichtige dabei ist, daß für Temperaturen unterhalb von 0,3 K die leichte He-3-reiche Phase aus nahezu reinem He-3 besteht, während die schwere He-4-reiche Phase auch bei $T = 0$ K noch 6% He-3 enthält. Diese Phase zeichnet sich durch einige thermodynamische Besonderheiten aus. Ihre He-4-Komponente befindet sich im suprafluiden Zustand und hat deshalb unterhalb von 0,3 K eine verschwindend kleine Entropie. Träger der gesamten thermischen Energie und Entropie dieser Phase sind die gelösten He-3-Atome. Das He-4 übernimmt für die He-3-Atome die Funktion eines Vakuums, man spricht vom „Quasivakuum", in dem sich die He-3-Atome als „Quasigas" bewegen. Der Übergang von He-3 aus der oberen Phase in die untere entspricht einer Quasiverdampfung von He-3 in das Quasivakuum He-4. Der Vorgang ist mit einer Lösungswärme verknüpft, die bei Temperaturen unterhalb von 0,05 K quadratisch mit der Temperatur abnimmt. Geschieht diese Quasiverdampfung bei thermischer Isolation, so wird die dazu notwendige Lösungswärme der Flüssigkeit entzogen und diese kühlt sich ab. Man muß dafür sorgen, daß ständig He-3-Atome aus der unteren Phase entfernt werden, so daß ständig He-3-Atome aus der oberen Phase über die Phasengrenze unter Aufnahme von Lösungswärme übergehen. Dieser Prozeß läßt sich leicht kontinuierlich gestalten und hat zum Bau von He-3-He-4-Kältemaschinen geführt, mit denen Temperaturen bis etwa 5 mK erzeugt und für unbegrenzte Zeit aufrechterhalten werden können.

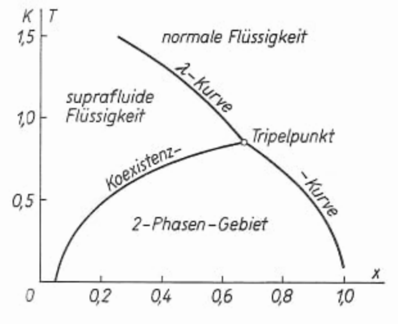


Abb. 6.3.6-3 Phasendiagramm von He-3-He-4-Lösungen; x bedeutet die He-3-Konzentration

6.4. Fester Aggregatzustand

Für feste Stoffe sind zwei Grenzzustände der Ordnung denkbar. Beim ideal isotropen Zustand sind die einzelnen diskreten Partikel, Atome, Ionen, Moleküle, vollkommen ungeordnet, d. h., sie zeigen eine statistische Verteilung. Bei einem idealen Kristall nehmen die Bausteine dagegen feste Ortslagen ein. Dieser Zustand läßt sich jedoch nicht streng realisieren, da Kristalle selbst am absoluten Nullpunkt noch einen bestimmten Betrag an *Nullpunktsenergie* besitzen, der sich in geringfügigen, bei höheren Temperaturen zunehmenden Schwingungen der Atome um ihre Gleichgewichtslagen äußert.
Die überwiegende Zahl der in der Natur vorkommenden festen Stoffe tritt in kristalliner Form auf. Für die Darlegung des Aufbaus der Festkörper und der Symmetrieeigenschaften von Kristallen werden die Teilchenschwingungen zunächst außer acht gelassen und nur die Schwerpunktslagen der Bausteine betrachtet (Tafel 36).

6.4.1. Gitteraufbau, Gittertypen, Kristallsysteme, Kristallklassen, Raumgruppen

6.4.1.1. Gitteraufbau. Die für einen Idealkristall charakteristische dreidimensional periodische Anordnung der Bausteine läßt sich veranschaulichen, indem wir von einem Teilchen, symbolisch dargestellt durch einen Punkt, ausgehen und durch periodische Translation in den drei Raumkoordinaten ein Gitter aufbauen. Dieses *räumliche Gitter* kann man sich auch in Teilschritten entstanden denken: Betrachtet man zunächst den Verschiebungsvorgang in einer Richtung, indem man das Ausgangsatom um einen Translationsvektor $\boldsymbol{a}$ verschiebt, so gelangt man zu der in Abb. 6.4.1-1a dargestellten periodischen Punktreihe, die man als *eindimensionales Gitter* bezeichnet. Der Betrag von $\boldsymbol{a}$, $a = a_0$, wird *Periode* (Identitätsperiode, Identitätsabstand) genannt. Läßt man auf das Ausgangsatom in einer von der ersten abweichenden Richtung eine zweite Translation $\boldsymbol{b}$ einwirken, so gelangt man zu einem *zweidimensionalen Gitter*. Es ist aus identischen Punkten aufgebaut, die folglich untereinander austauschbar sind, ohne daß sich die Struktur ändert (Abb. 6.4.1-1b). Schließlich gelangt man durch Translation des Ausgangsatoms in der dritten Raumkoordinate zum dreidimensionalen oder *Raumgitter* (Abb. 6.4.1-1c). Das durch die drei Vektoren $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ gebildete Parallelepiped wird *Elementarzelle* genannt und ist in Abb. 6.4.1-1c stark ausgezogen. Durch periodische Aneinanderreihung vieler Elementarzellen gelangt man zum makroskopischen Kristall.

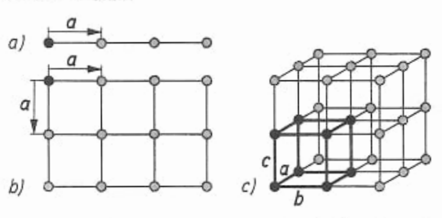


Abb. 6.4.1-1 a) eindimensionales Gitter (*Punktreihe*), b) zweidimensionales Gitter (*Punktnetz*), c) räumliches Translationsgitter

Im allgemeinen Fall bilden die Translationsrichtungen einen Winkel miteinander, der von 90° verschieden ist ($\alpha \neq \beta \neq \gamma$), und auch die Translationsbeträge sind in den verschiedenen Richtungen unterschiedlich ($a_0 \neq b_0 \neq c_0$). Das Gitter verhält sich somit anisotrop. Da das Gitter andererseits aus identischen Punkten bzw. Elementarzellen besteht, ist es zugleich homogen. Die Eigenschaft der *Anisotropie*, d. h. der grundsätzlichen Abhängigkeit des Verhaltens von der Richtung, wird bereits vollständig vom Begriff des aus identischen Punkten bestehenden Raumgitters erfaßt.

6.4.1.2. Kristallsysteme und Gittertypen. Es läßt sich zeigen, daß nicht beliebig viele Punktgitter, auch Translations- oder *Bravaisgitter* genannt, existieren, sondern daß ihre Zahl auf 14 begrenzt ist. Variiert man die Winkel α, β und γ zwischen den Achsen eines Translationsgitters und die Beträge a_0, b_0 und c_0 in den drei Rich-

Tab. 6.4.1-2 Die sieben Kristallsysteme

Kristallsystem	Achsenlängen	Winkel zwischen den Achsen
triklin	$a_0 \neq b_0 \neq c_0 \neq a_0$	$\alpha \neq \beta \neq \gamma \neq \alpha;\ \alpha, \beta, \gamma \neq 90°$
monoklin	$a_0 \neq b_0 \neq c_0 \neq a_0$	$\alpha = \gamma = 90°, \beta \neq 90°$
rhombisch	$a_0 \neq b_0 \neq c_0 \neq a_0$	$\alpha = \beta = \gamma = 90°$
hexagonal und trigonal	$a_0 = b_0 \neq c_0$	$\alpha = \beta = 90°, \gamma = 120°\ (60°)$
rhomboedrisch	$a_0 = b_0 = c_0$	$\alpha = \beta = \gamma \neq 90° \neq 60° \neq 109° 28' 16''$
tetragonal	$a_0 = b_0 \neq c_0$	$\alpha = \beta = \gamma = 90°$
kubisch	$a_0 = b_0 = c_0$	$\alpha = \beta = \gamma = 90°$

tungen unter Beachtung der Symmetrie, gelangt man zu sieben *Kristallsystemen* (Tab. 6.4.1-2).

Zu jedem Kristallsystem gehören ein oder mehrere Translationsgitter. Sie sind in den Tafeln 70/71 zusammengestellt. Jedes System enthält ein sogenanntes primitives oder einfaches Gitter, bei dem nur die Eckpunkte der Elementarzelle besetzt sind (*P-Gitter*); der Name rührt daher, daß jedes der 8 Teilchen nur zu einem Achtel der jeweiligen Elementarzelle angehört, so daß sich mit 8/8 nur eine einfache Besetzung ergibt. Darüber hinaus gibt es z. B. im rhombischen Kristallsystem ein basiszentriertes zweifach primitives *C-Gitter* (8/8 + 2/2), ein raum- oder innenzentriertes, ebenfalls zweifach primitives *I-Gitter* (8/8 + 1/1) und ein vierfach primitives flächenzentriertes *F-Gitter* (8/8 + 6/2).

Jeder Kristallstruktur liegt eines der 14 Translationsgitter zugrunde. Aus der geforderten Identität aller Gitterpunkte im Bravaisgitter folgt zwangsläufig, daß nur Stoffe, die aus *einer* Atomsorte bestehen, Translationsgitter aufbauen können. Eine Ausnahme bilden *Substitutionsmischkristalle*, in denen *A*-Atome durch *B*-Atome statistisch ersetzt sind, so daß sie strukturell als äquivalent betrachtet werden können.

Die Struktur jeder noch so komplizierten Verbindung läßt sich dann durch Ineinanderstellen verschiedener Translationsgitter aufbauen. Zum Beispiel besteht das in Tafel 71 dargestellte NaCl-Gitter, in dem sich Natrium- und Chloridionen abwechseln, aus den zwei achsenparallel verschobenen kubisch flächenzentrierten Translationsgittern von Na^+ und Cl^- (s. a. 6.4.1.5.).

Ganz allgemein legt man die Koordinaten des Translationsgitters so fest, daß sie den ausgezeichneten Richtungen des Kristalls entsprechen. Die Elementarzelle des kubisch kristallisierenden Natriumchlorids wird man also so wählen, daß ihre Kanten äußerlich erkennbaren Würfelrichtungen parallel verlaufen. Der Translationsbetrag der Teilgitter ist identisch mit der Gitterkonstante a_0, bei Steinsalz NaCl 563,8 pm, ist also nicht der Abstand r_0 zwischen Na^+ und Cl^-.

6.4.1.3. Kristallographische Symbolik

6.4.1.3.1. *Millersche Indizes.* Es ist ein auffallendes, wenn auch nicht notwendiges Merkmal vieler Kristalle, daß sie Polyeder bilden, die von ebenen Flächen begrenzt werden. Eine solche Kristallfläche schneidet im allgemeinen alle drei Achsen des zugrunde gelegten Koordinatensystems (Abb. 6.4.1-3); diese Achsenabschnitte werden benutzt, um die Lage der Kristallfläche zu beschreiben. Dazu werden die Achsenabschnitte aber nicht in derselben Einheit in den drei Richtungen gemessen, sondern in den jeweiligen Translationsabständen a_0, b_0 und c_0 oder in ganzzahligen Vielfachen davon, ma_0, nb_0 und pc_0; dabei bedeuten m, n und p ganze Zahlen.

In der Periodizität des Gitters liegt begründet, daß jeder Fläche im Kristall eine unendliche Schar identischer, parallel verlaufender *Netzebenen* (↑ 6.4.2.3.1.) zugeordnet ist, die ein konstantes Verhältnis der Achsenabschnitte aufweisen. Die Millerschen Indizes, das Symbol einer Netzebene, sind nun durch das Verhältnis der *reziproken* Achsenabschnitte definiert: $h:k:l = 1/m : 1/n : 1/p$. Die sich als Millersche Indizes ergebenden kleinen ganzen Zahlen werden in runde Klammern gesetzt: (hkl). Die in Abb. 6.4.1-3 gezeigte Fläche würde somit das Symbol (111) erhalten, da die Achsenabschnitte gerade $1a_0$, $1b_0$ und $1c_0$ betragen. Sind die Achsenabschnitte einer Fläche z. B. $2a_0$, $3b_0$ und $4c_0$, so ergibt sich als zugehöriges Flächensymbol gemäß $h:k:l = 1/2 : 1/3 : 1/4 = 6:4:3$ der Ausdruck (643).

Die Abschnitte, die eine Fläche auf den drei Achsen markieren, müssen nach dem Gesetz der rationalen Achsenabschnitte oder *Rationalitätsgesetz* zu den drei Längen a_0, b_0 und c_0 der Achsen in einem ganzzahligen Verhältnis stehen. Liegt die Fläche einer Achse parallel, so wird der Achsenabschnitt ∞ und der entsprechende Millersche Index somit 0. Die sechs Flächen eines Würfels schneiden jeweils nur eine Koordinatenachse und verlaufen zu den beiden anderen parallel. Von den sechs Flächen verlaufen außerdem jeweils zwei zueinander parallel. Drei Flächen tragen daher die Symbole $(\bar{1}00)$, $(0\bar{1}0)$ und $(00\bar{1})$, während die Gegenflächen mit einem negativen Vorzeichen versehen werden: (100), (010), (001). In Abb. 6.4.1-4 sind die Millerschen Indizes für verschiedene Flächen angegeben. Am Beispiel einer zweidimensionalen Darstellung (Abb. 6.4.1-5) ist erkennbar, daß die Indizes der einzelnen Flächenlagen um so kleiner sind, je dichter die betreffenden Netzebenen mit Gitterbausteinen besetzt sind. Niedrig indizierte Flächen haben also im allgemeinen eine hohe Besetzungsdichte.

Eine Besonderheit in der Indizierung nimmt das hexagonale System (Tafeln 70/71) ein.

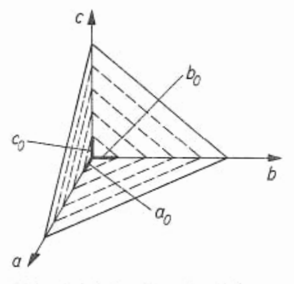


Abb. 6.4.1-3 Zur Indizierung einer Kristallfläche

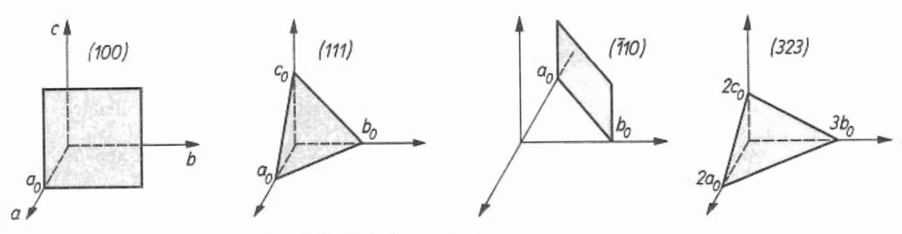


Abb. 6.4.1-4 Millersche Indizes verschiedener Flächen

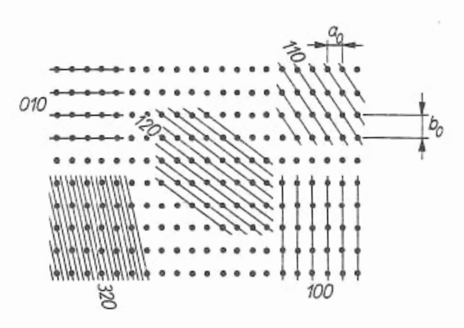


Abb. 6.4.1-5 Zusammenhang zwischen Millerschen Indizes und Besetzungsdichte einzelner Netzebenen

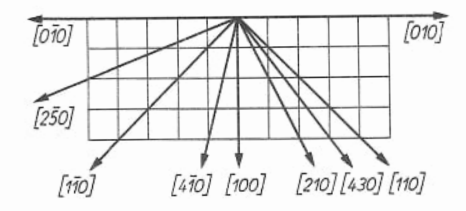


Abb. 6.4.1-6 Richtungssymbole für ein zweidimensionales Gitter in der (001)-Ebene

Hier benutzt man häufig kein dreiachsiges Koordinatensystem, sondern entsprechend der Symmetrie senkrecht zur sechszähligen Hauptachse c drei gleichwertige Achsen a_1, a_2 und a_3. Zur Kennzeichnung einer Fläche sind daher vier Indizes (*hkil*) notwendig, wobei allerdings die ersten drei nicht unabhängig voneinander sind. Es gilt die Beziehung $h + k + i = 0$. Die Basisfläche eines hexagonalen Kristalls trägt somit das Symbol (0001).

6.4.1.3.2. *Richtungssymbole.* Zur Unterscheidung von den Flächensymbolen (*hkl*) verwendet man für die Richtungssymbole [*uvw*] eckige Klammern. Sie dienen der Kennzeichnung einer beliebigen kristallographischen Richtung. In diesem Falle legt man das Verhältnis der Achsenabschnitte selbst zugrunde und nicht wie bei den Flächensymbolen das reziproke.

Die Richtung ist eindeutig bestimmt, wenn irgendein Punkt auf der Geraden bekannt ist. Bei Verschiebung des Punkts entlang der Geraden bleibt das Verhältnis der Koordinaten stets konstant. Zur Erläuterung der *Richtungssymbole* betrachten wir ein zweidimensionales Gitter in der (001)-Ebene (Abb. 6.4.1-6), wodurch der dritte Richtungsindex Null wird.

Jede Gerade kann parallel so verschoben werden, daß sie durch den Koordinatenanfangspunkt geht. Richtungen, die parallel zur a-Achse verlaufen, bekommen das Symbol [100], solche parallel zur b-Achse [010] und parallel zur c-Achse [001]. Die Gegenrichtung wird auch hier durch ein negatives Zeichen über dem Index angedeutet. Im hexagonalen System entspricht der c-Achse das Richtungssymbol [0001].

6.4.1.4. Kristallklassen und Symmetrie von Kristallen. Innere und äußere Symmetrie von Kristallen hängen eng zusammen. Die äußere, makroskopische Symmetrie wird durch bestimmte *Symmetrieelemente* beschrieben sowie durch die zugehörigen *Symmetrieoperationen*, die jeweils eine Wiederholung des Motivs bewirken. Am Kristall äußern sich als Symmetrieoperationen die *Drehung*, *Spiegelung*, *Inversion* sowie die *Inversionsdrehung* bzw. *Drehspiegelung*. Ihnen liegen als Symmetrieelemente zugrunde: *Drehachse* C_n, *Spiegelebene* m, *Inversionszentrum* Z und *Inversionsdrehachse* X bzw. *Drehspiegelachse* S_n (↑ 5.2.1.1.).

Je nach der Anzahl n der gleichwertigen Lagen, die während einer vollen Drehung bei Kristallen möglich sind, nennt man die Drehachsen oder *Gyren* [*gyros* griech., Wendung] 1-, 2-, 3-, 4- und 6-zählig. Die zugehörigen Symmetrieoperationen sind Drehungen um 360, 180, 120, 90 bzw. 60 Grad.

Symmetrie- oder Spiegelebenen m [*miroir* frz., Spiegel] teilen Kristalle in Hälften, die sich wie Bild und Spiegelbild verhalten. Durch die Spiegelungsoperation an den Symmetrieebenen wird jeweils Deckungsgleichheit erzeugt.

Ein Symmetrie- oder Inversionszentrum Z ist ein Punkt im Kristall, an dem durch die *Inversion*, die Spiegelung an einem Punkt, eine parallel liegende, gleichartige Gegenfläche erzeugt wird.

Tab. 6.4.1-7 Die 32 Kristallklassen

Kristall-system	Nr.	Klassennamen GROTH (1921)	Symbol SCHOEN-FLIES	HERMANN-MAUGUIN	Symmetrieelemente	
triklin	1	pedial	C_1	1	–	–
	2	pinakoidal	C_i	$\bar{1}$	–	Z
monoklin	3	sphenoidisch	C_2	2	1 ◊	–
	4	domatisch	C_s	*m*	1 *m*	–
	5	prismatisch	C_{2h}	2/*m*	1(♦ + *m*)	Z
rhombisch	6	rh.-disphenoidisch	D_2	222	1 ♦ + 1 ♦ + 1 ♦	–
	7	rh.-pyramidal	C_{2v}	*mm*2	1 ◊ + 1 *m* + 1 *m*	–
	8	rh.-dipyramidal	D_{2h}	*mmm*	1(♦ + *m*) + 1(♦ + *m*) + 1(♦ + *m*)	Z
tetragonal	9	t.-pyramidal	C_4	4	1 ▣	–
	10	t.-disphenoidisch	S_4	$\bar{4}$	1 ⊟	–
	11	t.-dipyramidal	C_{4h}	4/*m*	1(■ + *m*)	Z
	12	t.-trapezoedrisch	D_4	422	1 ■ + 2 ♦ + 2 ♦	–
	13	dit.-pyramidal	C_{4v}	4*mm*	1 ▣ + 2*m* + 2*m*	–
	14	t.-skalenoedrisch	D_{2d}	$\bar{4}2m$	1 ⊟ + 2 ♦ + 2*m*	–
	15	dit.-dipyramidal	D_{4h}	4/*mmm*	1(■ + *m*) + 2(♦ + *m*) + 2(♦ + *m*)	Z
trigonal	16	tr.-pyramidal	C_3	3	1 △	–
	17	rhomboedrisch	C_{3i}	$\bar{3}$	1 ◬	Z
	18	tr.-trapezoedrisch	D_3	32	1 ▲ + 3 ◊	–
	19	ditr.-pyramidal	C_{3v}	3*m*	1 △ + *m*	–
	20	ditr.-skalenoedrisch	D_{3d}	$\bar{3}m$	1 ◬ + 3(♦ + *m*)	Z
hexagonal	21	h.-pyramidal	C_6	6	1 ⬡	–
	22	tr.-dipyramidal	C_{3h}	$\bar{6}$	1 ⌬ ≡ ▲ + *m*	–
	23	h.-dipyramidal	C_{6h}	6/*m*	1(⬢ + *m*)	Z
	24	h.-trapezoedrisch	D_6	622	1 ⬢ + 3 ♦ + 3 ♦	–
	25	dih.-pyramidal	C_{6v}	6*mm*	1 ⬡ + 3*m* + 3*m*	–
	26	ditr.-dipyramidal	D_{3h}	$\bar{6}m2$	1 ⌬ (≡ ▲ + *m*) + 3 ◊ + *m*	–
	27	dih.-dipyramidal	D_{6h}	6/*mmm*	1(⬢ + *m*) + 3(♦ + *m*) + 3(♦ + *m*)	Z
kubisch	28	tetraedrisch-pentagon-dodekaedrisch	*T*	23	3 ♦ + 4 △	–
	29	disdodekaedrisch	T_h	*m*3	3(♦ + *m*) + 4 ◬	Z
	30	pentagonikositetra-edrisch	*O*	432	3 ■ + 4 ▲ + 6 ♦	–
	31	hexakistetraedrisch	T_d	$\bar{4}3m$	3 ⊟ + 4 △ + 6*m*	–
	32	hexakisoktaedrisch	O_h	*m*3*m*	3(■ + *m*) + 4 ◬ + 6(♦ + *m*)	Z

Zeichenerklärung

♦ ▲ ■ ⬢ 2-, 3-, 4-, 6-zählige Drehachsen C_n
◊ △ ▣ ⬡ 2-, 3-, 4-, 6-zählige polare Drehachsen
◬ ⊟ ⌬ 3-, 4-, 6-zählige Inversionsdrehachsen $\bar{X}$
m Symmetrieebene
Z Symmetriezentrum

Inversionsdrehachsen $\bar{X}$ stellen eine Kombination aus Drehachse C_n und Inversionszentrum Z dar. Die Deckoperation besteht folglich aus einer Drehung mit anschließender Inversion. Die ebenfalls mögliche Kombination von Drehachsen C_n mit Spiegelebenen m zu den Drehspiegelachsen oder Gyroiden S_n führt nicht zu zusätzlichen Symmetriemöglichkeiten. Beispielsweise ist eine dreizählige Drehspiegelachse S_3 identisch mit einer sechszähligen Inversionsdrehachse $\bar{6}$.

Da manche Symmetrieelemente einander zwangsläufig bedingen, ergeben sich nur 32 verschiedene Möglichkeiten einer Kombination der makroskopischen Symmetrieelemente miteinander; alle Kristalle lassen sich entsprechend ihrer äußeren Symmetrie einer dieser Möglichkeiten zuordnen, also einer von 32 *Kristallklassen*. Die alternative Bezeichnung *Punktgruppen* rührt daher, daß bei der Ausführung der entsprechenden Symmetrieoperationen wenigstens 1 Punkt des Kristalls ortsfest bleibt. In Übereinstimmung mit den jeweils vorhandenen Symmetrieelementen lassen sich die 32 Kristallklassen den 7 Kristallsystemen zuordnen. Wie Tab. 6.4.1-7 zeigt, tritt die niedrigste Symmetrie bei triklin-pedialen, die höchste bei hexakisoktaedrischen Kristallen auf. Die in der Tab. genannten *polaren Achsen* treten nur bei Kristallklassen ohne Symmetriezentrum auf; den verschiedenen Enden einer polaren Achse kommen verschiedene Eigenschaften zu, z. B. durch die Besetzung mit Ionen entgegengesetzten Vorzeichens.

6.4.1.5. Raumgruppen. Die innere, mikroskopische Symmetrie der Kristalle läßt sich aus der äußeren, makroskopischen dadurch herleiten, daß als weitere Symmetrieoperation die *Translation* berücksichtigt wird. Aus einem Symmetrieelement wird dadurch eine Parallelschar unendlich vieler Symmetrieelemente erzeugt.

Die Translation kann auf die Kombination der Symmetrieelemente, wie sie in einer Punktgruppe jeweils verwirklicht ist, als Ganzes angewendet oder mit einzelnen Punktsymmetrieoperationen gekoppelt werden. Als neue Mikrosymmetrieoperationen ergeben sich durch Kombination von Translation und Drehung die *Schraubung*, von Translation und Spiegelung die *Gleitspiegelung*. Die entsprechenden Mikrosymmetrieelemente sind *Schraubenachse* und *Gleitspiegelebene*. Aus Abb. 6.4.1-8a, b ist ersichtlich, daß durch viermalige Drehung um je 90° und gleichzeitige Verschiebung um $a/4$ ein Atom wieder seine identische Lage erreicht; bei Abb. 6.4.1-8c wird durch einfache Spiegelung Identität erreicht, während bei 6.4.1-8d das Atom 1 erst nach zweimaliger Spiegelung und Translation um je $a/2$ in die identische Lage 1′ gelangt.

Die Anwendung von Translation, Schraubung und Gleitspiegelung auf die 32 Kristallklassen führt dann zu 230 unterschiedlichen Symmetriekombinationen, die als *Raumgruppen* bezeichnet werden. Die tatsächlichen *Raumgitter* kristalliner Festkörper lassen sich damit als Bravaisgitter darstellen, die so ineinandergestellt sind, daß sie die Bedingungen der Raumgruppensymmetrie erfüllen. Das ist bei jeder der 230 Raumgruppen für unendlich viele Strukturen möglich.

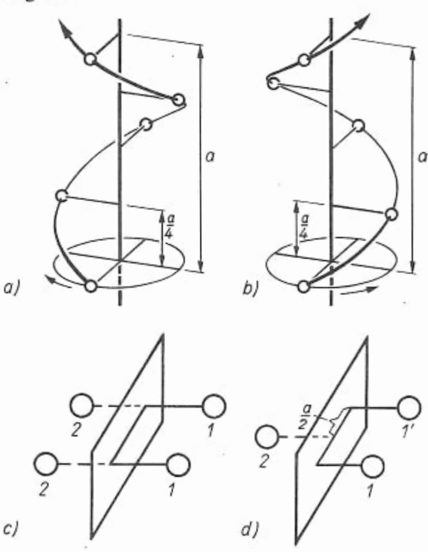


Abb. 6.4.1-8 a) vierzählige linke, b) vierzählige rechte Schraubenachse, c) Spiegelebene, d) Gleitspiegelebene

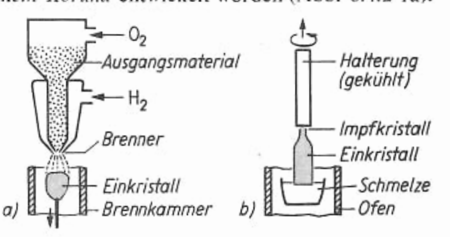
6.4.2. Kristalline Stoffe

Die Möglichkeiten zur kristallographisch eindeutigen Beschreibung kristalliner Festkörper wurden unter 6.4.1. behandelt. Da aber selbst bei weitgehend identischen Kristallstrukturen verschiedener Festkörper große Unterschiede in deren physikalischen und chemischen Eigenschaften auftreten können, empfiehlt sich die Klassifizierung kristalliner Stoffe auch nach anderen Gesichtspunkten:

Die Einteilung nach der Art der Bindungskräfte zwischen den Gitterbausteinen (↑ 6.4.3.) erleichtert das Verständnis für die Strukturen der chemischen Elemente und der Verbindungen (↑ 6.4.4.). Die phänomenologische *Thermodynamik* (↑ 7.1.2.) abstrahiert zwar ihrem Wesen nach von Bindungs- und Strukturfragen, in der Abhängigkeit von Temperatur, Druck oder Zusammensetzung geben sich andererseits selbst geringfügige strukturelle Umwandlungen oder spezifische energetische Wechselwirkungen zu

erkennen. Insofern sind die *Zustandsdiagramme* unter Beteiligung fester Phasen (↑ 6.4.5.) für das Verständnis von Festkörperstabilität und -reaktivität von großem Wert.
Zustandsdiagramme gelten im allgemeinen für den Gleichgewichtszustand des betrachteten Systems. In neuerer Zeit gewinnen aber auch *nichtkristalline Stoffe* (↑ 6.4.6.) große Bedeutung, die aufgrund ihrer Herstellung stark vom Gleichgewichtszustand abweichen.
Zwischen diesen amorphen Materialien und den Idealkristallen stehen solche Festkörper, die noch Gitterperiodizität, aber schon deutliche Störungen in der Anordnung der Kristallbausteine aufweisen. Diese *Realstruktur von Kristallen* (↑ 6.4.7.) ist für viele Fragen der Festkörperaktivität von entscheidender Bedeutung. Die Zusammenhänge zwischen *Kristallsymmetrie* sowie *gitterphysikalischen Eigenschaften* (↑ 6.4.8., 8.1.) werden durch Abweichungen von der Idealstruktur teils sehr empfindlich gestört, teils nur unwesentlich beeinflußt.

6.4.2.1. Kristallzüchtung. *Einkristalle* werden meist aus der Schmelze des gewünschten Stoffs, aus dessen Lösung in einem geeigneten Lösungsmittel oder aus der Gasphase unter Umgehung des flüssigen Aggregatzustands abgeschieden. Ihre Herstellung aus kristallinen Festkörpern gleicher oder veränderter Zusammensetzung ist demgegenüber auf Ausnahmen beschränkt.
Die Technik der *Kristallzüchtung* hat durch den zunehmenden Bedarf an Reinststoffen in den letzten Jahrzehnten eine stürmische Entwicklung erfahren. Dabei haben die Verfahren zur Kristallisation aus der Schmelze die stärkste Verbreitung gefunden. Eine hohe Perfektion in der Ausbildung der Kristallstruktur wird erreicht, wenn nur ein wachstumsfähiger Keim vorliegt und der Kristall durch weitgehende Annäherung an die Gleichgewichtsbedingungen für den Phasenübergang hinreichend langsam wächst. Für jedes Verfahren zur Züchtung großer Einkristalle steht die Aufgabe, den durch die freiwerdende *Kristallisationsenthalpie* in der Probe auftretenden Temperaturgradienten möglichst weitgehend zu unterdrücken oder zumindest die dadurch entstandenen Spannungen durch eine Nachbehandlung abzubauen
Die *Verneuilmethode* ist für die Züchtung von *synthetischen Rubinen* oder *Saphiren* sowie von *reinem Korund* entwickelt worden (Abb. 6.4.2-1a). Das Ausgangsgemisch der entsprechenden Oxide, Aluminium- und Chromium(III)-oxid für Rubin, Eisen(III)-, Titanium(IV)- und Aluminiumoxid für Saphir, wird in feingepulverter Form durch einen Vibrator in eine Knallgasflamme gegeben und eingeschmolzen. Auf einer vertikal beweglichen Unterlage wächst der Einkristall in mehreren Stunden in Form einer Birne; nach sorgfältiger Temperung zur Entfernung von inneren wachstumsbedingten Spannungen ist ein Zerschneiden in die gewünschte Form, z. B. mittels einer Diamantsäge, möglich.
Beim Verfahren nach Bridgman und Stockbarger wird das Ausgangsmaterial in einem Gefäß eingeschmolzen, das am Boden in einer Spitze ausläuft. Bei langsamer Entfernung des Bodens aus der Zone maximaler Temperatur gelingt es, einen zuvor eingebrachten *Impfkristall* so wachsen zu lassen, daß er schließlich das gesamte Gefäß ausfüllt. Auf diese Weise werden viele Metalle, aber auch organische und anorganische Verbindungen hergestellt. Durch Arbeiten unter Inertgas lassen sich Störungen durch Reaktion mit Luftsauerstoff ausschließen, jedoch ist mit Verunreinigungen des Kristalls durch das Wandmaterial zu rechnen.
Der Kontakt des wachsenden Kristalls mit dem Tiegel wird bei dem weitverbreiteten *Czochralskiverfahren* unterbunden (↑ Abb. 6.4.2-1 b). Ein Impfkristall, gehaltert an einer vertikalen Dreh-Zug-Vorrichtung, wird mit der Oberfläche der Schmelze des jeweiligen Stoffs in Kontakt gebracht. Durch geeignete Temperaturführung läßt sich die Kristallisation gerade an der Phasengrenze Schmelze – Kristall weiterführen. Unter langsamer Rotation wird der wachsende Kristall aus der Schmelze gezogen; es entsteht ein Stab von kreisförmigem Querschnitt und einem Durchmesser bis etwa 100 mm, der in Scheiben der gewünschten Dicke zerteilt wird. Auf diese Weise lassen sich *Silicium*, *Germanium*, *Galliumarsenid* GaAs, *Galliumphosphid* GaP, *Indiumantimonid* InSb und *Rubine* herstellen.
Ein elegantes Prinzip zur tiegelfreien Herstellung von Einkristallen ist im *Schwebe-* oder *Fließzonenverfahren* (↑ Abb. 6.4.2-1c, Tafeln 72/73) verwirklicht. Hierbei wird ein Stab des polykristallinen Materials durch Induktionsheizung von einem Ende her aufgeschmolzen und durch Wandern des Temperaturfelds über



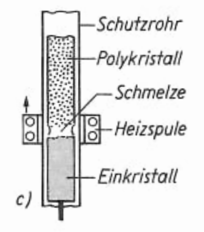


Abb. 6.4.2-1 Methoden der Einkristallzüchtung (schematisch); a) Verneuilverfahren, b) Czochralskiverfahren, c) Fließzonenverfahren

die Länge des Stabs zum langsamen Erstarren, nunmehr als Einkristall, gebracht. Der Zusammenhalt zwischen polykristallinem Ausgangsmaterial und wachsendem Einkristall wird durch die Oberflächenspannung der Schmelze gewährleistet. Die Fließzonentechnik ist für die Züchtung von Silicium- und Germaniumeinkristallen optimiert worden. Das Verfahren kann als sogenanntes *Zonenschmelzen* auch zur weiteren Reinigung von Einkristallen bzw. zur Vorreinigung eines für die Kristallzüchtung vorgesehenen Materials herangezogen werden. Man nutzt die Tatsache aus, daß Verunreinigungen sich meist nicht gleichmäßig auf flüssige und feste Phase verteilen. In der Regel ist die Konzentration der Verunreinigung in der schmelzflüssigen Phase höher als in der festen. Beim lokalen Aufschmelzen des Ausgangseinkristalls und bei seiner erneuten Erstarrung reichern sich also die Verunreinigungen in der wandernden Schmelzzone an. Durch mehrfache Wiederholung des Zonenschmelzens kann die *chemische Fehlordnung* (↑ 6.4.7.) um mehrere Zehnerpotenzen verringert werden. Diese Technik wird außer bei Halbleitern auch zur Herstellung ultrareiner, oft überraschend dehnbarer Metalle (↑ 6.4.8.) und solcher organischer Verbindungen eingesetzt, die wegen ihrer optischen Eigenschaften (s. a. 8.8.5.) von Interesse sind.

Bei der Kristallzucht aus Lösung besteht neben der Möglichkeit einer Kontamination mit Gefäßmaterial auch die des Einschlusses kleiner Lösungsmittelmengen. Für zahlreiche Substanzen technisch ausgereift ist die *Hydrothermalsynthese*, bei der die natürlichen Wachstumsbedingungen in wäßrigen Lösungen simuliert, durch geeignete hohe Temperaturen und Drücke aber stark beschleunigt werden. So gelingt z. B. die Züchtung von *Quarzkristallen*, da dann selbst für das in Wasser unter Normalbedingungen kaum lösliche Siliciumdioxid nennenswerte Konzentrationen in Lösung und relativ hohe Kristallisationsgeschwindigkeiten erreichbar werden.

Noch höhere Anforderungen an die experimentelle Beherrschung hoher Drücke und Temperaturen stellt die *Diamantsynthese* durch Umwandlung von Graphit bei Drücken bis zu 7500 MPa ($\approx$ 75000 atm) und Temperaturen von 3000 °C. Die entstehenden Kristalle sind wegen ihrer Kleinheit meist der industriellen Anwendung vorbehalten.

Die Züchtung gutausgebildeter Einkristalle direkt aus der Gasphase kann unterhalb des aus dem Phasendiagramm ersichtlichen *Tripelpunktdrucks* (↑ 6.4.5.) erfolgen. Die Verwendung von Impfkristallen ist auch hier zweckmäßig. Eine spontane Keimbildung wäre erst bei erheblicher Übersättigung des Dampfs (↑ 6.3.2.3.) wahrscheinlich, es käme dann aber zur Ausbildung vieler kleiner Kristalle, die sich wegen des relativ geringen Substanzangebots nur skelettartig in Richtung größter Kristallwachstumsgeschwindigkeit ausbilden könnten, wie es von Schneeflocken geläufig ist.

Zersetzt sich eine Verbindung beim Verdampfen, kann ihre einkristalline Abscheidung durch eine *Transportreaktion*, d. h. über Bildung und Zerfall flüchtiger Zwischenverbindungen, oft dennoch gewährleistet werden.

Für die Herstellung mikroelektronischer Schaltkreise auf Siliciumbasis ist ein Verfahren entwickelt worden, bei welchem durch *Epitaxie* (↑ 8.3.6.2.2.) auf anderweitig erhaltenen Einkristallen orientiert Silicium aus dem Dampf abgeschieden wird, z. B. durch Reaktion von *Siliciumtetrahalogeniden* mit Wasserstoff bzw. *Silan*, SiH_4. Eine heteroepitaktische Abscheidung einer Substanz auf einer anderen, aber von der Gittergeometrie her geeigneten kristallinen Unterlage spielt bei der Herstellung bestimmter $A^{III}B^{V}$*-Verbindungen* für die Optoelektronik (↑ 8.8.3.) eine Rolle.

Für die Herstellung von gewollt, d. h. durch *Dotierung* „verunreinigten" Einkristallen, wie sie z. B. von der Halbleitertechnik gefordert werden, bietet die Kristallzüchtung aus der Gasphase durch gleichzeitige thermische Zersetzung bzw. Reduktion flüchtiger Verbindungen der Haupt- und der Dotierungskomponente die Möglichkeit einer relativ leichten Einstellung der Zusammensetzung der wachsenden Einkristallschicht.

In jüngster Zeit ergeben sich durch die Kosmosforschung neue Möglichkeiten für die Züchtung besonders perfekter Einkristalle. Durch die in Orbitalstationen um mehrere Größenordnungen verminderte Gravitation sind Störungen des Kristallwachstums durch Dichtegradienten und Konvektionsströmungen in der Schmelze weitgehend ausgeschlossen.

6.4.2.2. Strukturanalyse kristalliner Körper. Licht- und vor allem Elektronenmikroskope gestatten zwar meist, Produkte als kristallin zu identifizieren, ihr Auflösungsvermögen genügt aber trotz großer Fortschritte der Höchstspannungselektronenmikroskopie noch nicht für eine direkte Abbildung der Anordnung der Gitterbausteine. Hierzu muß Strahlung einer Wellenlänge verwendet werden, die kleiner oder gleich dem durchschnittlichen Atomabstand im Gitter ist.

Der Nachweis der Beugung von Röntgenstrahlung an Kristallen durch Laue, Friedrich und Knipping veranlaßte W. L. Bragg und W. H. Bragg bald zu Versuchen, die erhaltenen *Beugungsdiagramme* zur Ermittlung der *Kristallstruktur* auszunutzen. Aus der Erkenntnis von De Broglie, daß bewegten Körpern eine Wellennatur zukomme (↑ 1.2.4.), ließ sich die Erwartung ableiten, daß Elektronen, Neutronen oder andere Partikel geeigneter kinetischer Ener-

gie ebenfalls an Kristallgittern gebeugt werden müßten.
Die simultane Untersuchung der Struktur mittels *Röntgen-*, *Elektronen-* und *Neutronenbeugung* ist besonders empfehlenswert bei nichtkristallinen festen Körpern. Dort werden anstelle scharfer Beugungsreflexe breite Banden in der sogenannten *Radialverteilungsfunktion* erhalten. Sie ließen sich bei Anwendung nur einer Beugungsmethode lediglich für solche Substanzen eindeutig auswerten, die sämtlich aus Atomen einer Art bestehen (s. a. 6.4.6.1.).
Wegen der unterschiedlichen Wechselwirkung der Atome mit den einzelnen Strahlungen ergänzen sich die genannten Methoden vorteilhaft auch bei der Untersuchung kristalliner Stoffe. So erlaubt die *Röntgenbeugung* beim Vorliegen leichter Atome, z. B. von Wasserstoff, in der betreffenden Substanz oft keine genaue Lokalisierung von deren Positionen; dagegen ist die Beugung von Neutronen auch an Wasserstoffatomen stark. Bei solchen Vergleichen ist allerdings zu bedenken, daß die Beugung von Elementarteilchen vor allem an der *Kristalloberfläche* erfolgt, die in Zusammensetzung und Struktur oft beträchtlich gegenüber dem Kristallinnern modifiziert ist.

6.4.2.2.1. *Röntgenbeugung; Braggsche Gleichung.* Betrachten wir Röntgenstrahlen der Wellenlänge λ, die auf eine Kristallfläche auftreffen, welche durch regelmäßig angeordnete, gleichsam in den Kreuzungspunkten eines Gitters lokalisierte Atome gebildet wird. Durch die Wechselwirkung des elektrischen Felds der Röntgenwellen mit den Elektronen in den Hüllen dieser Atome gehen von jedem Atom Kugelwellen einer Streustrahlung aus, die sich wegen ihrer Inkohärenz meist gegenseitig auslöschen. In Phase befinden sich solche – jeweils von benachbarten Punkten ausgehende – Wellen, deren Gangunterschied ein ganzzahliges Vielfaches der Wellenlänge λ, also $n\lambda$, beträgt. Für $n = 0$ liegt Beugung nullter Ordnung vor, d. h., die Strahlen bilden einen Beugungskegel, der sich in der Richtung der Primärstrahlen ausbreitet. Er ist zum Unterschied von den Beugungsmaxima höherer Ordnung unabhängig vom Abstand zwischen den Streuzentren und somit völlig strukturunspezifisch.
Die Röntgenbeugung am dreidimensionalen Kristall wurde verständlich, als W. L. Bragg die Möglichkeit erkannt hatte, die komplexen Beugungserscheinungen auf das einfache geometrische Problem scheinbarer Reflexionen von Röntgenstrahlen an Kristallflächen zurückzuführen. Hierzu betrachtet man eine Schar paralleler *Netzebenen* (Tafel 27) mit dem Abstand d voneinander, auf die ein Bündel Röntgenstrahlen der Wellenlänge λ unter dem Winkel Θ auftrifft (Abb. 6.4.2-2). Die Strahlen werden teils von der obersten, teils von der 2., 3., … n. Ebene reflektiert. Sie verlassen den Kristall dann in gleicher Phase, wenn die Differenzen in ihren Weglängen ganzzahlige Vielfache von λ sind. Zum Beispiel weisen die auf benachbarte Netzebenen auffallenden Strahlen, A und B bzw. B und C, jeweils einen Gangunterschied $n\lambda = 2x$ zueinander auf. Da nun $x = d \sin \Theta$, ergibt sich als *Braggsche Gleichung*

$$(1) \quad n\lambda = 2d \sin \Theta; \qquad n = 1, 2, 3, \ldots$$

In Anpassung an die Art des Probenmaterials, die verwendete Röntgenstrahlung und die zu erwartende Symmetrie des Kristalls wurden verschiedene Techniken zur Strukturanalyse entwickelt. Der Braggschen Reflexionsbedingung wird man in jedem Fall gerecht. Entweder hält man den Kristall und damit Θ fest und variiert λ, oder man benutzt monochromatische Röntgenstrahlung und verändert Θ, was durch Drehen des Einkristalls während der Aufnahme bzw. durch Verwendung von Kristallpulver gelingt.

6.4.2.2.2. *Einkristallmethoden der Röntgenbeugung.* Vollständige Strukturbestimmungen erfordern die Messung der Reflexionsintensitäten einer möglichst großen Zahl von Netzebenenscharen. Zur Untersuchung reichen *Einkristalle* mit Abmessungen von etwa 0,2 bis 0,5 mm in den drei Raumrichtungen völlig aus.
Besonders günstig zur Ermittlung der Kristallsymmetrie ist das *Laueverfahren*, das aus den Anfängen der Röntgenbeugung an Kristallen herrührt. Dabei wird ein feststehender Einkristall mit „weißem" Röntgenlicht durchstrahlt (Abb. 6.4.2-3a). Entsprechend der variablen Wellenlänge ist die Braggsche Gleichung [↑ 6.4.(1)] für zahlreiche Netzebenen erfüllt. Auf einen senkrecht zum Primärstrahl angebrachten Planfilm kommen die Reflexe in derselben Anordnung um den Primärfleck zu liegen, in der die Netzebenen übereinstimmender d-Werte zum Röntgenstrahl orientiert sind. So weist das Beugungsbild eine sechszählige Symmetrieachse auf, wenn der Röntgenstrahl parallel zu einer sechszähligen Drehachse des Kristalls einfällt.
Bei der Auswertung der *Lauediagramme* ist allerdings zu beachten, daß statt der 32 auf-

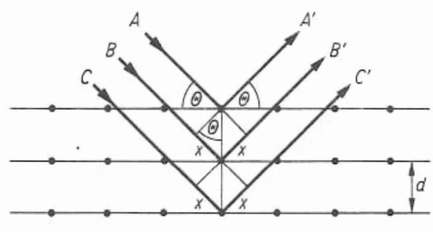


Abb. 6.4.2-2 Ableitung der Braggschen Gleichung

grund der Punktsymmetrieelemente unterscheidbaren Kristallklassen (↑ 6.4.1.) nur diejenigen 11 beobachtet werden, die ein Symmetriezentrum aufweisen. Durch die Spezifik der Methode werden auch für Kristallklassen geringerer Symmetrie zentrosymmetrische Reflexe erhalten, die Symmetrie wird also in diesen Fällen auf eine der 11 *Lauegruppen* erhöht. Das ist ein Beispiel für das allgemeine Prinzip, wonach die Symmetrieelemente einer Kristalleigenschaft wohl die des Kristalls enthalten, aber nicht vollständig mit ihnen zusammenfallen müssen.

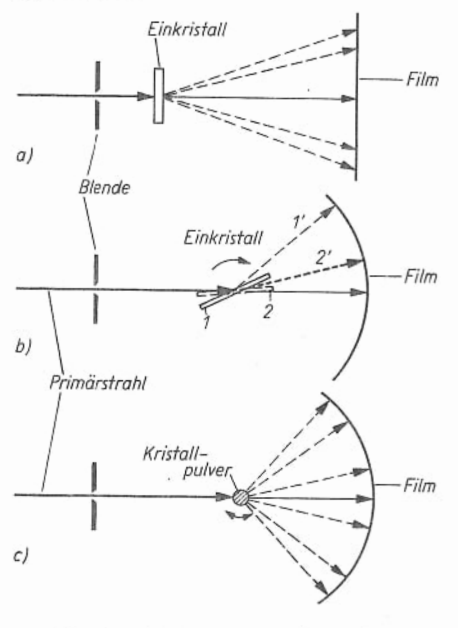


Abb. 6.4.2-3 Methoden der Röntgenbeugung an Kristallen; a) Laue-, b) Drehkristall-, c) Debye-Scherrer-Verfahren

Bei der *Drehkristallmethode* (↑ Abb. 6.4.2-3 b) wird der Kristall so justiert, daß er senkrecht zur Richtung des monochromatischen Röntgenstrahls um eine Achse gedreht werden kann. Der zur Registrierung verwendete Film liegt auf der Innenseite einer zylindrischen Kamera. Die zur Drehachse parallel liegenden Scharen von Netzebenen ergeben die sogenannte nullte Schicht- oder Äquatorlinie, auf der sich entsprechend den jeweiligen *Glanzwinkeln* Θ die Reflexe des Kristalls anordnen. Die übrigen Netzebenenscharen bilden entsprechende, ober- und unterhalb des Äquators liegende 1., 2., ... *n*. Schichtlinien. Aus deren Abständen sind Schichtlinienwinkel, Identitätsperioden im Kristall, bei geeigneter Orientierung unmittelbar die jeweiligen Gitterkonstanten der Elementarzelle erhältlich.

Varianten der Drehkristallmethode arbeiten zur genaueren Erfassung einzelner Schichtlinien mit einer synchron zur Kristalldrehung ablaufenden Filmbewegung. Neuere Geräte ermitteln die Intensitäten aller auftretenden Einzelreflexe mit Zählrohren. Durch hohen Automatisierungsgrad und Kopplung mit leistungsfähigen Computern gelingt die Aufklärung auch komplizierter Strukturen in der relativ kurzen Zeit von Tagen bis Wochen (Tafeln 32/33).

6.4.2.2.3. *Pulvermethoden der Röntgenbeugung.* Läßt man monochromatische Strahlung auf ein sehr feines Pulver regellos zueinander orientierter Kristallite fallen, entstehen auch hier scharfe Beugungsdiagramme, da stets für eine gewisse Zahl von Kristalliten die für den jeweiligen Glanzwinkel Θ erforderliche Orientierung zum Primärstrahl gegeben ist. Das *Debye-Scherrer-Verfahren* arbeitet mit fotografischer Registrierung auf einem zylindrisch um das Präparat angeordneten Film (↑ Abb. 6.4.2-3c). Die geforderte Konstanz von λ wird durch die verwendete einfache Filterung nur annähernd gewährleistet, was sich in einer gewissen Linienverbreiterung äußert. Beim *Guinierverfahren* wird dagegen durch Einsatz eines *Kristallmonochromators* strenge Monochromasie des Primärstrahls und hohe Auflösung erreicht. Auch *Pulverdiffraktometer*, die an ebenen Reflexionspräparaten die Beugungsintensität in Abhängigkeit vom Ablenkungswinkel mittels eines Zählrohrs registrieren, besitzen eine hohe Meßgenauigkeit.

6.4.2.2.4. *Elektronen- und Neutronenbeugung.* Elektronen werden vom Kristallgitter wesentlich stärker beeinflußt als Röntgenstrahlen. Sehr dünne Schichten können in Transmission, dickere Proben unter streifendem Einfall des Elektronenstrahls untersucht werden, allerdings auch hier nur mit Eindringtiefen im Nanometerbereich. Die *Elektronenbeugung* ist also besonders zur Untersuchung von Oberflächen geeignet, z. B. zum Nachweis der Bildung neuer Phasen. Es treten deutliche Linien schon bei viel geringeren Anteilen einer Phase auf, als sie für ein auswertbares Röntgendiffraktogramm erforderlich wären. Günstig ist auch, daß die Linienverbreiterung der Reflexe erst bei sehr viel kleineren Kristalliten auftritt als bei der Röntgenbeugung. Das rührt daher, daß sich schon mit üblichen Beschleunigungsspannungen um 50 kV Elektronenwellenlängen um 5 pm, also beträchtlich kleinere als bei üblicher Röntgenstrahlung ($Cu_{K\alpha}$ = 154 pm), ergeben. Der weitere Vorteil der Elektronenbeugung, mit wesentlich kürzeren Belichtungszeiten als die Röntgenbeugung auszukommen, ist durch die modernen Entwicklungen der Röntgentechnik relativiert worden.

Wegen der erheblichen Wechselwirkung der Elektronenstrahlen mit dem Feld der Atomkerne sind auch leichte Atome, deren Hüllen durch Röntgenstrahlung nur wenig beeinflußt werden, gut zu lokalisieren; z. B. läßt sich der Wasserstoffeinbau in Halbleiterschichten quantitativ verfolgen. Noch leistungsfähiger sind hier allerdings Beugungsuntersuchungen mit Elektronen sehr geringer Energie, *LEED* [low energy electron diffraction engl.], von etwa 10 bis 100 eV, da hierbei nur die ersten Atomlagen der Oberfläche erfaßt werden, so daß auch Aufwachsmechanismen, z. B. bei epitaktischer Abscheidung, studiert werden können.
Mit *thermischen Neutronen*, die eine Geschwindigkeit bis 10^6 cm s^{-1} und dementsprechend Wellenlängen im 100-pm-Bereich aufweisen, ergeben sich ebenfalls Besonderheiten für die Strukturuntersuchung. Die Streuung erfolgt im wesentlichen an den Atomkernen und ist wegen der Elektroneutralität der Neutronen generell schwach. Der an sich sehr hohe Aufwand für die Erzeugung intensiver Neutronenströme im Kernreaktor und ihre Monochromatisierung bleibt in Grenzen, da die hohe Nachweisempfindlichkeit der Zählrohre eine genaue Absolutmessung selbst schwacher Reflexintensitäten für die jeweiligen Netzebenen ermöglicht.
Die Amplitude der kohärenten Streuung der Neutronen kann sich schon bei verschiedenen Isotopen desselben Elements meßbar unterscheiden, andererseits bleiben die Unterschiede zwischen so verschiedenen Elementen wie Uranium und Beryllium so gering, daß die Atomlagen jeweils beider Elemente ermittelt werden können.
Besonders wertvoll ist die Möglichkeit der Bestimmung der *magnetischen Struktur* von Verbindungen. Da das Neutron ein magnetisches Moment aufweist, überlagert sich der Wechselwirkung von Neutron und Kern die Wechselwirkung des Neutrons mit den Magnetmomenten der Atome des Kristallgitters. Durch das Studium der *Neutronenbeugung* in Abhängigkeit von der Temperatur lassen sich *ferro-*, *antiferro-*, *ferri-* oder *paramagnetischer Zustand* voneinander abgrenzen. Feinheiten der Lückenbesetzung der *Spinellstruktur* (↑ 7.3.3.2.3.) von *Ferriten* sind ebenso zu erkennen wie die Ausbildung von *Überstrukturphasen* bei Legierungen. Durch Neutronenbeugung unter kleinen Winkeln werden auch Besonderheiten der Domänenstruktur (↑ 8.6.4.4.) und selbst Auswirkungen von mechanischer Bearbeitung oder Temperbehandlung auf Metalle erkennbar.
In neuerer Zeit ergänzt die *Protonenbeugung* die übrigen Beugungsverfahren gelegentlich dadurch, daß sie besonders auf wesentliche Züge der Realstruktur anspricht (Tafeln 34 und 35).

6.4.3. Bindungstypen bei Festkörpern und energetische Kristallchemie

Für die energetische Beschreibung von Festkörpern gibt es hauptsächlich zwei Wege. Die *Festkörperphysik* faßt den Kristall vorzugsweise als Ensemble von Kernen, Elektronen und Quasiteilchen auf, während von der *Kristallchemie* der Kristall als Raumgitter betrachtet wird, in dessen ausgezeichneten Lagen sich die Atome der jeweiligen chemischen Elemente befinden. Dementsprechend ist die Kristallchemie bemüht, den Atomen auch im Kristall individuelle Eigenschaften zuzuschreiben, wie Radien, Ladungen, Koordinationszahlen und Elektronegativitäten, und den energetischen Zusammenhalt des Gitters über die Wechselwirkungen zwischen den Atomen, Ionen bzw. Molekülen, also dem jeweiligen Bindungstyp entsprechend, zu erfassen. Die übliche, wenn auch nicht die einzig mögliche Unterscheidung ist die in *Ionen-*, *Metall-*, *Atom-* und *van-der-Waals-Gitter*.
Die *chemische Bindung in Kristallen* ist oft beträchtlich modifiziert gegenüber der in isolierten Molekülen, da dann das Verhalten der Valenzelektronen durch das periodische Potentialfeld des Kristalls und damit durch die Raumanordnung der Teilchen im Gitter und die Stöchiometrie der Verbindungen bestimmt wird. Der Vorzug der Anschaulichkeit der kristallchemischen Betrachtungsweise gegenüber der festkörperphysikalischen wird dadurch erkauft, daß die Annäherung an die Realität oft nicht sehr weitgehend ist. Das trifft besonders zu auf die, aus den Anfängen der Kristallchemie überkommene, übertrieben weitgehende Anwendung des vereinfachten *Ionenmodells* auf den Bau anorganischer Kristalle. Bei den Metallkristallen hat sich eine kristallchemische Beschreibung der Bindung im Gitter durch PAULING, auf Grundlage der Valenzbindungstheorie, gegenüber dem Bändermodell (↑ 8.2.) der Festkörperphysik gar nicht erst durchsetzen können. Die energetische Kristallchemie interessiert sich für die *Gitterenergie* unterschiedlich gebundener Festkörper, da sie gesetzmäßig mit verschiedenen anderen Kristalleigenschaften verknüpft ist. Die theoretische Berechnung der Gitterenergien erfolgt über geeignete Ansätze für die Wechselwirkungspotentiale; die Überprüfung der Werte wird dadurch möglich, daß man mit experimentell erhaltenen thermochemischen Daten *Kreisprozesse* zur indirekten Ermittlung der Gitterenergie formuliert.
Die Gitterenergie U_G ist definiert als die Energie, die frei wird, wenn ein Mol isolierter Bausteine zum Kristallgitter zusammentritt. Bei Ionenkristallen sind die isolierten Bausteine die unendlich weit voneinander entfernten freien Ionen, bei überwiegend kovalent gebundenen Stoffen die einzelnen Atome. Die Gitterenergie

solcher Atomgitter ist somit bis auf das Vorzeichen identisch mit der sogenannten *Atomisierungsenthalpie*, bei den Metallen zugleich auch mit deren *Sublimationsenthalpie*, da Metalle in der Regel ohnehin einatomige Dämpfe bilden. Bei Feststoffen, in denen die Gitterbausteine lediglich durch die van-der-Waals-Kräfte zusammengehalten werden, ist der Bezugszustand für die Gitterenergie unterschiedlich, von Atomen bei den Edelgasen, zweiatomigen Molekülen bei Wasserstoff und den Halogenen bis zu Molekülen beachtlicher Größe bei höheren Kohlenwasserstoffen.

Man könnte meinen, daß bei einem Festkörper die Struktur mit der größten Gitterenergie immer zugleich die stabilste Teilchenanordnung darstelle. Für die Stabilität eines Zustands ist allerdings nicht dessen Energie U bzw. Enthalpie H, sondern die freie Energie F bzw. freie Enthalpie G entscheidend (↑ 6.1.2.3.2.). Für die zu deren Berechnung benötigten Entropien S der einzelnen Gitter unterschiedlicher Symmetrie existieren jedoch noch keine zuverlässigen Ansätze. Die Modifikation eines Stoffs mit der höchsten Gitterenergie kann folglich nicht von vornherein auch als die thermodynamisch stabilste angesprochen werden.

6.4.3.1. Gitterenergieberechnungen bei Ionenkristallen. Heteropolare oder *Ionengitter* werden als aus Ionen aufgebaut betrachtet, vorzugsweise aus solchen mit Edelgaskonfiguration der Valenzelektronen; die Stabilität dieser Anordnung führt zu einer geringeren potentiellen Energie der Ionen im Gitter als bei der Kombination freier Atome. Aus den *Ionisierungsenergien* und *Elektronenaffinitäten* der chemischen Elemente (↑ S. 730/731) folgt, daß die Neigung zur Bildung von Ionen mit wachsendem Abstand der Atome von der Gruppe der Edelgase abnehmen wird, d. h. also von den *Alkalimetallhalogeniden* $A^{I}B^{VII}$ über die *Erdalkalimetallchalkogenide* $A^{II}B^{VI}$, die $A^{III}B^{V}$-*Verbindungen* bis zu den Elementen der IV. Hauptgruppe. Nur die ersten beiden Substanzklassen können als ionisch aufgebaut angesehen werden. Die $A^{III}B^{V}$-Halbleiter *Galliumarsenid*, *Indiumantimonid* usw. lassen sich besser mit den Kovalenz- oder Atomgittern von *Silicium* und *Germanium* vergleichen.

Als typischer Vertreter der Ionenkristalle gilt *Natriumchlorid*; nach der klassischen Paulingschen Abschätzung ergibt sich für die Na-Cl-Einzelbindung aus der Elektronegativität der Atome eine Polarität von etwa 65%. Das NaCl-Gitter ist kristallchemisch aufzufassen als regelmäßige Anordnung von sechs Chloridionen um ein Natriumion und umgekehrt, entsprechend dem Ineinanderschieben von zwei kubisch flächenzentrierten Bravaisgittern (↑ 6.4.1.2.) um eine halbe Würfelkante. Dabei überwiegen die Anziehungskräfte zwischen den ungleichartigen benachbarten Teilchen die Abstoßungskräfte zwischen den gleichartigen, weiter voneinander entfernten.

Der Hauptanteil der Gitterenergie U_G wird durch die *Coulombwechselwirkung* der Ionen untereinander bewirkt, die im 1. Term der Gleichung dargestellt ist

$$(1) \quad U_G = -\frac{N_A e^2 A_M}{4\pi\varepsilon_0 r_0} + \frac{N_A B}{r_0^n} + U_{\mathrm{vdW}} + U_0$$

(2. Term Abstoßungspotential nach BORN; 3. Term van-der-Waals-Energie; 4. Term Nullpunktsenergie). Hierin bedeuten: N_A Avogadrokonstante, e Elementarladung, A_M Madelungkonstante, ε_0 elektrische Feldkonstante, r_0 Gleichgewichtsabstand benachbarter ungleich geladener Ionen im Kristall, B positive Konstante, n Bornexponent ($n \gg 1$, etwa 9 bis 12). Zur Berechnung des Coulombterms stellt man sich die Ionen als starre Kugeln vor, deren Ladung allerdings punktförmig im Mittelpunkt konzentriert gedacht wird. Ein Ion wird als Zentrum ausgewählt und seine elektrostatische Wechselwirkung mit allen anderen Ionen betrachtet. Hierbei ist eine Summation über die Anziehungs- und Abstoßungsterme bis zu recht großen Abständen von diesem Zentrum notwendig, da die Wechselwirkung mit wachsendem r nur mit der ersten Potenz des Teilchenabstands r abnimmt.

Gegenüber der Bildung ausschließlich isolierter Na^+Cl^--Ionenpaare ergibt sich durch die kollektive Wechselwirkung der Ionen über den gesamten Kristall hinweg eine energetische Stabilisierung, die durch die *Madelungkonstante* A_M erfaßt wird. Für die NaCl-Struktur nimmt sie den Wert $A_{NaCl} = 1{,}7476$ an, weitere Werte sind in Tab. 6.4.3-1 enthalten. Bei Kenntnis des Gleichgewichtsabstands r_0 zwischen benachbarten ungleich geladenen Ionen kann mit Hilfe dieser Konstanten für alle Substanzen des jeweiligen Strukturtyps der Coulombanteil an der Gitterenergie berechnet werden. Für Stoffe, die nicht nur einwertige Ionen und/oder nicht mehr die einfache Stöchiometrie 1 : 1 aufweisen, nehmen die so errechneten Energien hohe Werte an, die über die Stabilisierung der jeweiligen Struktur ein irreales Bild vermitteln könnten, z. B. für Magnesiumoxid MgO 4625 kJmol^{-1}, für Thoriumdioxid ThO_2 23100 kJmol^{-1}, für NaCl dagegen nur 862 kJmol^{-1}. Durch verschiedene Normierungen lassen sich weitere Madelungkonstanten ableiten, die klarere Aussagen darüber gestatten, welche Struktur für ein beliebiges Ion energetisch effektiver wäre (↑ Tab. 6.4.3-1). Die Coulombenergie würde den Wert $-\infty$ erreichen, wenn nicht bei entsprechend kleinen Ionenabständen, $r \leqq r_0$, eine starke Abstoßung sowohl zwischen den äußeren Elektronenhüllen der Ionen als auch zwischen

den positiv geladenen Atomrümpfen zustande-käme. In der hohen Potenz von r_0 im Ausdruck für das Abstoßungspotential 6.4(1) kommt gleichsam das Wirken des Pauliprinzips zum Ausdruck, nach dem eine gegenseitige Durchdringung vollbesetzter Elektronenschalen verboten ist, da sie zu einer Mehrfachbesetzung der jeweiligen, durch die vier Quantenzahlen n, m, l und s bestimmten Zustände führen würde (↑ 4.2.2.3.). Die hohe Potenz von r_0 bewirkt, daß es meist genügt, nur die nächsten Nachbarn, also die 1. Koordinationssphäre, in die Rechnung einzubeziehen. Der Abstoßungsterm nimmt etwa 10% des Zahlenwerts der Coulombenergie an. Er steht in Zusammenhang mit der *Kompressibilität*, läßt sich also näherungsweise auch experimentell ermitteln.

Ein dritter Beitrag zur Gitterenergie resultiert aus der van-der-Waals-, speziell der Londonwechselwirkung (↑ 6.4.3.2.), die unabhängig vom Bindungstyp stets vorhanden ist, nur über sehr kurze Distanzen wirkt und der Polarisierbarkeit beider Teilchensorten proportional ist. Meist weit unter dem Betrag der Coulombenergie liegend, erreicht dieser Term doch bei den höheren Caesium- und Thalliumhalogeniden den Betrag der Bornschen Abstoßungsenergie.

Einen vierten Beitrag zur Gitterenergie liefern die Schwingungen, die die Gitterbausteine noch am absoluten Nullpunkt ausüben. Die damit verbundene gewisse Unbestimmtheit in Teilchenlage und -impuls läßt sich aufgrund der *Heisenbergschen Unschärferelation* (↑ 2.5.1.7.) verstehen. Die Beträge der *Nullpunktsenergie* liegen meist deutlich unter 1% der Gesamtgitterenergie und können oft vernachlässigt werden.

Andererseits sind die Gesamtenergiedifferenzen zwischen möglichen Alternativstrukturen gewöhnlich gleichfalls nur gering. Entsprechend den Madelungfaktoren A_{CsCl} und A_{NaCl} läge die Coulombenergie von Natriumchlorid in einer CsCl-Stuktur nur knapp 1% über der für den NaCl-Typ, Konstanz des interionischen Abstands r_0 vorausgesetzt. Berücksichtigt man, daß die drei weiteren Terme, die zur Gitterenergie U_G beitragen, so genau gar nicht berechnet werden können, werden die Unsicherheiten bei der Voraussage von Modifikationswechseln verständlich.

Die durch theoretische Ansätze ermittelten Gitterenergien können durch sogenannte „experimentelle" Werte überprüft werden, die auf indirekte Weise aus thermochemischen Daten zugänglich sind. Für Ionenkristalle wird meist der *Born-Habersche Kreisprozeß* zugrunde gelegt. Er stellt ein Beispiel für den Hessschen Satz (↑ 6.1.2.2.4.) dar und beschreibt den für die Gitterenergie U_G von Ionenkristallen wesentlichen Vorgang der Bildung des Gitters aus freien Ionen über eine Reihe thermochemisch zugänglicher Teilreaktionen. Er wird am Beispiel von *Natriumchlorid* in Abb. 6.4.3-2 veranschaulicht. Die einzelnen Größen bedeuten hierbei: $\Delta_B H^\ominus$ molare Standardbildungsenthalpie von NaCl, $\Delta_S H^\ominus$ molare Standardsublimationsenthalpie von Natrium, $\Delta_D H^\ominus$ molare

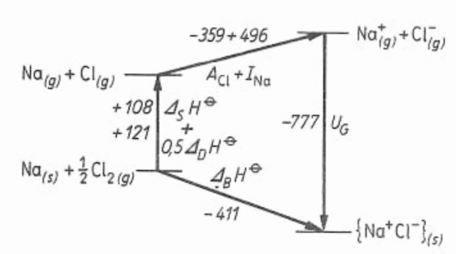


Abb. 6.4.3-2 Born-Haberscher Kreisprozeß für Natriumchlorid; Werte in kJ mol^{-1}

Tab. 6.4.3-1 Madelungkonstanten A_M und A_{red} für verschiedene Strukturtypen; die nur noch geometrieabhängigen reduzierten Werte ergeben sich zu $A_{red} = 2\,A_M/z_+ \cdot z_- \cdot m$; z_+; z_- Ionenladungen, m Zahl der Teilchen je Formeleinheit

Strukturtyp	A_M	z_+	z_-	A_{red}
isolierte Ionenpaare AB	1,000	1	1	1,000
Cuprit (Cu_2O)	4,4425	1	2	1,481
Rutil (TiO_2)	4,816	2	1	
	19,083	4	2	1,605
Korund (α-Al_2O_3)	24,242	3	2	1,616
Zinkblende	1,638	1	1	1,638
(Sphalerit ZnS)	6,552	2	2	
Wurtzit (ZnS)	1,6413	1	1	1,641
	6,563	2	2	
Fluorit (CaF_2)	5,039	2	1	1,639
Perowskit ($CaTiO_3$)	24,755	2, 4	2	1,658
Steinsalz (NaCl)	1,7476	1	1	1,748
Caesiumchlorid (CsCl)	1,7627	1	1	1,763

Standarddissoziationsenthalpie von Chlor, I_{Na} molare Ionisierungsenergie von Natrium, A_{Cl} molare Elektronenaffinität von Chlor, U_G Gitterenergie oder genauer die molare Standardbildungsenthalpie des Ionengitters aus den freien Ionen. Es ergibt sich $U_{G(NaCl)} = -I_{Na} - A_{Cl} - \Delta_S H^{\ominus}_{Na} - \frac{1}{2}\Delta_D H^{\ominus}_{Cl_2} + \Delta_B H^{\ominus} = (-496 + 359 - 108 - 121 - 411)\ \mathrm{kJmol^{-1}} = -777\ \mathrm{kJ \cdot mol^{-1}}$. Aus Gleichgewichtsmessungen in der Dampfphase wurde als direkter experimenteller Wert $U_{G(NaCl)} = -756\ \mathrm{kJmol^{-1}}$ erhalten; die vorgenannten theoretischen Ansätze liefern Werte $U_{G(NaCl)} = -750$ bis $-770\ \mathrm{kJmol^{-1}}$.

Bei Erdalkalimetallhalogeniden und -chalkogeniden treten, besonders bei Kombinationen schwererer Elemente, weit größere Abweichungen zwischen den errechneten und experimentell bestimmten Energien der Ionengitter auf. Für diese Substanzen erweist sich das dem Coulombansatz zugrunde liegende Ionenmodell als schon zu stark vereinfacht.

Eine kugelsymmetrische Ladungsverteilung der Valenzelektronen ist streng genommen nur bei freien edelgasähnlichen Ionen gegeben. Auf Ionen im Kristall wirkt ein elektrisches Feld, das die Punktsymmetrie des von ihnen besetzten Gitterkomplexes aufweist, also nicht Kugelsymmetrie erreicht. Somit geht von den umgebenden Ionen eine polarisierende Wirkung auf das betrachtete Ion aus; dieses wirkt seinerseits ebenfalls polarisierend, so daß auch die gegenseitige *Polarisierbarkeit* der Ionen in Rechnung zu setzen ist.

Die Polarisierbarkeit von Atomionen nimmt mit deren Ordnungszahl und der Größe der negativen Ionenladung zu. Die *Polarisationswirkung* kann grob über den Quotienten z/r_a^2 (mit z Ionenladung, r_a Ionenradius) abgeschätzt werden, nur grob deshalb, weil sowohl für die Ladung z als auch für die Radien r_a stark idealisierende Annahmen gemacht werden müssen (s. a. 6.4.4.). Jedenfalls wirken Teilchen mit kleinem Raumbedarf und hoher effektiver Ladung stark polarisierend, so daß bei Verwendung der klassischen Ionenradien (↑ 6.4.4.2.) die Kationen meist die stärker polarisierenden, die Anionen die stärker polarisierbaren Teilchen sind. Mit wachsender Anzahl nächster Nachbarn eines Ions erhöht sich die Umgebungssymmetrie, so daß die für Gitterenergierechnungen bei Ionengittern postulierte Kugelsymmetrie der Ionen am ehesten bei den höheren Koordinationszahlen 8 und 6 erfüllt sein sollte. Bei nur vier oder gar drei nächsten Nachbarn ist dagegen die Gitterpolarisierung so stark, daß man die Ionenladung nicht mehr als im Zentrum starrer, inkompressibler Kugeln befindlich betrachten kann. Der Nutzen der *Radienquotientenregel* (↑ 6.4.4.3.) zur Festlegung des Gittertyps von Ionenkristallen wird dadurch stark eingeschränkt.

6.4.3.2. Berechnung der Energie von van-der-Waals-Gittern. Organische und andere Substanzen ohne nennenswerte Bindungspolarität werden im Kristall durch dieselben *van-der-Waals-Kräfte* zusammengehalten, die auch die Abweichungen von der Idealität im Verhalten von Gasen bewirken (s. a. 6.2.1.). Die van-der-Waals-Kräfte spielen für die Entscheidung über die Kristallstruktur kaum eine Rolle, da sie räumlich ungerichtet und außerdem nur sehr schwach sind. Die Struktur wird vielmehr wesentlich durch die Struktur der Teilchen selbst bestimmt. Bei Energieberechnungen für *van-der-Waals-Gitter* ist zu unterscheiden zwischen *Orientierungs-* (KEESOM), *Induktions-* (DEBYE) und *Dispersionsenergien* (LONDON).

Haben Moleküle permanente Dipolmomente, orientieren sie sich so, daß jeweils das positive Dipolende des einen Moleküls das negative eines anderen anzieht. Eine zusätzliche Wechselwirkung ergibt sich aus der Anziehung zwischen einem *permanenten Dipol* und den durch ihn in den Nachbarteilchen induzierten Momenten. Aber selbst völlig unpolare Moleküle und Atome, wie Anthracen oder Argon, üben Anziehungskräfte aufeinander aus, die letztlich zur Verflüssigung bzw. Kristallisation führen. Hierfür ist die London- oder Dispersionsenergie verantwortlich zu machen, die dadurch zustandekommt, daß auch bei Molekülen ohne permanentes Dipolmoment nicht in jedem Augenblick die Ladungsschwerpunkte der negativen Elektronen oder der positiven Kerne zusammenfallen können, sich also *momentane Dipole* ausbilden. Jeder momentane Dipol induziert dann im Nachbarmolekül einen weiteren.

Alle drei Beiträge zur van-der-Waals-Gitterenergie sind umgekehrt proportional der 6. Potenz des Abstands zwischen den Teilchen, was die äußerst geringe Reichweite der Wechselwirkungen erklärt. Für genauere Rechnungen ist auch bei van-der-Waals-Kristallen ein Abstoßungsterm zu berücksichtigen, der eine beliebige Annäherung der Partikeln verhindert. Die Gitterenergien bei van-der-Waals-Kristallen liegen etwa bei 10 bis 50 $\mathrm{kJmol^{-1}}$, also beträchtlich unter denen bei Ionen- oder Atomgittern. Das erklärt auch deren vergleichweise niedrige Schmelz- bzw. Sublimationstemperaturen.

6.4.3.3. Gitterenergieberechnungen bei kovalent gebundenen Kristallen. Die Berechnung der einzelnen zur Gitterenergie beitragenden Terme ist erheblich komplizierter als bei Ionen- und van-der-Waals-Kristallen, weil sich für die Energie der einzelnen Bindungen zwischen den Partnern wegen der komplexen gegenseitigen Beeinflussung vieler Teilchen weder experimentell noch quantenchemisch allgemeingültige Werte ableiten lassen. Dazu kommt, daß bei gleicher Valenzelektronenkonfiguration der Atome nicht notwendig gleichartige kovalente Bindungen an-

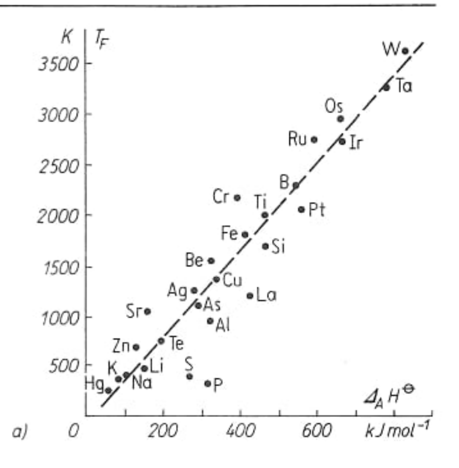
zunehmen sind. Die generelle Tendenz zu abnehmender Hybridisierung von s- und p-Orbitalen mit wachsender Ordnungszahl führt z. B. in der IV. Hauptgruppe dazu, daß der Prototyp eines Atomgitters, das *Diamantgitter* mit der sp^3-Hybridisierung der Valenzelektronen, nicht durchgängig bei allen Homologen verwirklicht ist. Der Übergang zur metallischen Bindung unter Erhöhung der Zahl der nächsten Nachbarn ist im weißen *Zinn* und im *Blei* nicht zu übersehen. Gitterenergieberechnungen, die diese Metallisierungstendenz erfassen und Modifikationswechsel vorauszusagen gestatten, stehen noch aus.

Leitet man *mittlere molare Bindungsenergien* aus einem thermochemischen Kreisprozeß her, begibt man sich faktisch von vornherein auf die Stufe der „experimentellen" Bestimmung der Gitterenergie, hat also keinen Gewinn gegenüber der unmittelbaren Verwendung der Atomisierungsenergie als einer summarischen Größe, die keinen Bezug zu Konstitution und Konfiguration der Atome hat.

Die Gitterenergie kovalent gebundener Stoffe mit dreidimensionaler Vernetzung der Struktureinheiten ist bis auf das Vorzeichen identisch mit deren molarer Standardatomisierungsenthalpie $\Delta_A H^\ominus$. Sie ist die zum Überführen des aus 1 mol Atomen zusammengesetzten Gitters in den Zustand der isolierten Atome benötigte Energie. Man kann sie für eine beliebig zusammengesetzte Verbindung $A_z B_y C_x$ ... über einen Kreisprozeß berechnen, der die molaren Standardatomisierungsenthalpien $\Delta_A H_i^\ominus$ der Komponenten i ($i = A, B, C, \ldots$), deren Molenbrüche x_i ($x_A, x_B, x_C, \ldots$) und die molare Standardbildungsenthalpie $\Delta_B H^\ominus_{A_z B_y C_x}$ ... der entsprechenden Verbindung verknüpft:

$$\Delta_A H^\ominus_{A_z B_y C_x}\ldots = -\Delta_B H^\ominus_{A_z B_y C_x}\ldots + \sum x_i \cdot \Delta_A H_i^\ominus .$$

Für eine Auswahl chemischer Elemente sind die Werte $\Delta_A H_i^\ominus$ in Tab. 6.4.3-3 angegeben; statt ihrer die molaren Standardsublimationsenthalpien $\Delta_S H_i^\ominus$ zu verwenden, kann nicht generell empfohlen werden, da die Sublimation einiger Elementarsubstanzen, z. B. von Schwefel, Selen, Tellur, Phosphor, Arsen und Antimon, nicht zwangsläufig schon zu den Atomen führt, wie es bei den Metallen noch die Regel ist.



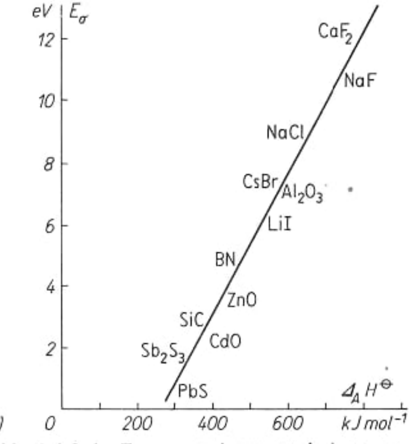


Abb. 6.4.3-4 Zusammenhang zwischen molarer Standardatomisierungsenthalpie $\Delta_A H^\ominus$ ausgewählter Stoffe und *a*) deren Schmelztemperaturen T_F sowie *b*) deren Aktivierungsenergien E_σ der elektrischen Leitfähigkeit (nach Urusov 1975)

Tab. 6.4.3-3 Molare Standardatomisierungsenthalpien $\Delta_A H_i^\ominus$ ausgewählter Elemente i; 298 K, alle Angaben in kJ mol⁻¹, Meßunsicherheit im allgemeinen 0,5 bis 2 kJ mol⁻¹

H	218,0	Pb	195
Li	161	P	334
Na	108	As	303
K	89	Sb	269
Cs	77	Bi	210
Be	325	O	249,2
Mg	147	S	279
Ca	179	Se	227
Sr	165	Te	196,5
Ba	184	F	79,5
B	561	Cl	121,3
Al	330	Br	111,8
Ga	272	I	106,8
Tl	180	Cu	337
C	715	Ag	285
Si	451	Zn	130,5
Ge	375	Cd	112
Sn	303	Hg	61,5

Für isostrukturelle, dreidimensional vernetzte Stoffe ergeben sich oft gesetzmäßige Zusammenhänge zwischen der Energie des Atomgitters und den physikalischen Eigenschaften des jeweiligen Stoffs. Abb. 6.4.3-4 läßt sich so interpretieren, daß das *Schmelzen* ebenso wie der *Ladungsträgertransport* vom Valenz- in das Leitungsband (↑ 8.2.1.4.2.) eine ähnliche, wenn auch nicht so weitgehende Aktivierung oder gar Zerstörung von Bindungen voraussetzt wie der Übergang des Atomgitters in den Zustand einatomiger Dämpfe. Für viele weitere, auch stärker ionisch gebundene Stoffe lassen sich angenähert lineare Zusammenhänge zwischen Atomisierungsenthalpie und Schmelztemperatur, *Härte*, *Kompressibilität* sowie Aktivierungsenergie der elektrischen Leitfähigkeit aufstellen.

Bei Gittern, deren Atome nur in ein oder zwei Raumrichtungen durch kovalente Bindungen zusammengehalten werden, ist es oft nicht mehr sinnvoll, Gitterenergie und Atomisierungsenthalpie gleichzusetzen. Da zwischen den Ketten oder Schichten oft nur die viel schwächeren van-der-Waals-Kräfte wirken, erfordert z. B. das Schmelzen solcher Stoffe weitaus geringere Energien als deren Atomisierung. Viskositätsmessungen ergaben z. B., daß die Aktivierungsenergie des viskosen Fließens bei dem kettenstrukturierten Selen mit $E_\eta = 40\ \text{kJmol}^{-1}$ weit unter der mittleren molaren Bindungsenergie von Selenbindungen, $E_{Se-Se} = 225\ \text{kJmol}^{-1}$, liegt, während für das dreidimensional vernetzte Germaniumdiselenid $GeSe_2$ mit den Energien $E_\eta = 240\ \text{kJ mol}^{-1}$ und $E_{Ge-Se} = 230\ \text{kJmol}^{-1}$ weitgehend übereinstimmende Werte erhalten wurden.

Realistische Gitterenergieberechnungen für Selen, Graphit und andere ein- bzw. zweidimensional vernetzte Stoffe müßten also die Zwischenstellung der entsprechenden Gitter zwischen dem van-der-Waals- und dem Atomgitter berücksichtigen.

6.4.4. Geometrische Kristallchemie

Zur kristallchemischen Beschreibung von Strukturen genügt im allgemeinen die Kenntnis der Abstände der Schwerpunkte benachbarter Atome voneinander und der Bindungswinkel zwischen ihnen. Damit ist die Symmetrie der Anordnung festgelegt. Eine Fernordnung der Baugruppen wird zwar als selbstverständlich vorausgesetzt, ist aber für die meisten Argumentationen der geometrischen Kristallchemie nicht von Belang. Es geht dort vielmehr darum, aus Größe und Ladung der Bausteine und aus der Eigensymmetrie der Bindungen die Kristallstrukturen der Elemente und ihrer Verbindungen zu verstehen, um zusammen mit der energetischen Kristallchemie (↑ 6.4.3.) die Stabilität und Reaktivität von Festkörpern bewerten zu können.

6.4.4.1. Strukturtypen bei den chemischen Elementen. Für gleichartige Teilchen mit kugelsymmetrischer Ladungsverteilung ist die Ausbildung *dichtester Kugelpackungen* (Abb. 6.4.4-1) am wahrscheinlichsten. Hier hat jedes Atom 12 nächste Nachbarn; 6 umgeben das Zentralatom in einer Schicht *A*, je 3 weitere berühren es von oben bzw. unten. Für sie gibt es zwei Varianten der Einbettung in drei der je-

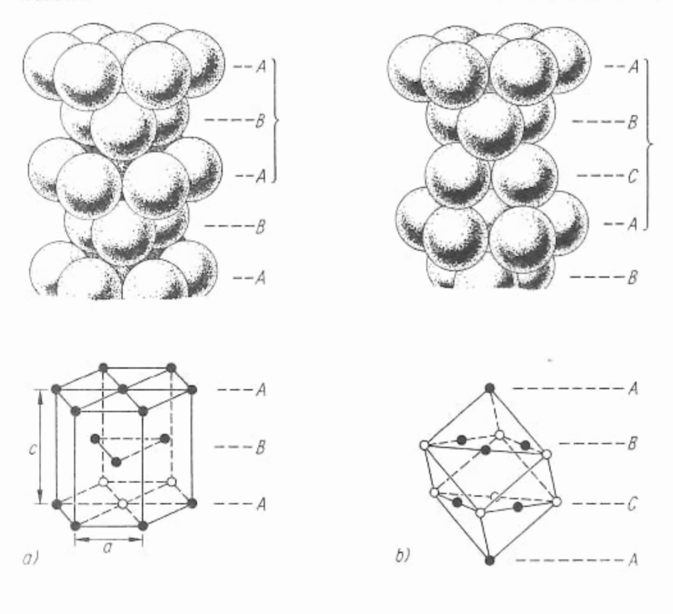


Abb. 6.4.4-1 Schichtfolge bei dichtesten Kugelpackungen und Zusammenhang mit den Elementarzellen; a) hexagonal, b) kubisch-flächenzentriert (die schwarzen Kugeln entsprechen den Teilchen, die in den durch die Klammer markierten Schichten gut erkennbar sind)

weils sechs Lücken um ein Atom der Ausgangsschicht *A*:
Bei der *hexagonal dichtesten Packung* werden in darüber- und darunterliegender Schicht die gleichen drei Positionen besetzt, so daß die Schichtenfolge *BABABA* ... resultiert. Die *kubisch dichteste Packung* entsteht bei der alternativen Besetzung dieser Plätze; es ergibt sich die Schichtenfolge *CABCABCAB* ... In ihr erkennt man die vierfach primitive, kubisch flächenzentrierte Elementarzelle, die dem Bravaisgitter *F* entspricht (↑ Abb. 6.4.4-1), wenn man sich vergegenwärtigt, daß die Senkrechte auf die Schichten *CABCAB* ... mit der Raumdiagonalen des Elementarwürfels zusammenfällt (↑ Abb. 6.4.4-1). Bei der hexagonal dichtesten Packung wird die Senkrechte zu *BABA* ... identisch mit der sechszähligen Hauptachse *c* des Bravaisgitters; für das Achsenverhältnis $c/a = 1{,}633$ haben alle 12 Nachbarn gleiche Abstände vom Zentralatom.
Dichteste Kugelpackungen treten bei den *Edelgasen* sowie bei der Mehrzahl der *Haupt-* und *Nebengruppenmetalle* auf. Es existieren auch Übergänge zwischen den Grenzfällen kubisch und hexagonal dichtester Packungen, und zwar dadurch, daß sich kubische Schichtpakete mit hexagonalen abwechseln, regelmäßig z. B. bei *Lanthan*, *α-Praseodymium* und *Neodymium*, oder unregelmäßig wie bei *Cobalt*.
Die *Alkalimetalle* bilden kubisch raumzentrierte, also *I*-Gitter (↑ Tafel 70/71). Bei diesem Strukturtyp mit der Koordinationszahl 8 geht die Raumerfüllung, die sich auf der Grundlage des vereinfachten Modells sich berührender starrer Kugeln berechnen läßt, auf 68 % zurück; bei den dichtesten Packungen beträgt sie 74 %.
Bei den Elementen der III. bis VII. Hauptgruppe fällt das Überwiegen der Atombindung bei den leichteren Vertretern, oft im Verein mit der Ausbildung von Hybridorbitalen, dagegen die abnehmende Hybridisierungstendenz und das allmähliche Hervortreten des Metallcharakters beim Übergang zu den schwereren Elementen auf (s. a. 6.4.3.3.). So hat *Bor* drei Modifikationen komplizierter Struktur mit ungleichwertigen Atompositionen im Gitter, während schon bei *Aluminium* mit der Koordinationszahl 12 das metallische Verhalten dominiert.
Für das Verständnis der Atomgitter bei den Elementen der VII. bis IV. Hauptgruppe ist die *(8-N)-Regel* nützlich. Ist *N* die Nummer der Gruppe im Periodensystem, gibt die Differenz $8 - N$ die Bindigkeit der jeweiligen Atome an.
VII. Hauptgruppe: Die einbindigen Halogenatome treten wie Wasserstoff zu zweiatomigen Molekülen X_2 und diese zu van-der-Waals-Gittern (↑ 6.4.3.2.) zusammen.
VI. Hauptgruppe: Chalkogenatome verwirklichen ihre Zweibindigkeit unterschiedlich. *Sauerstoff* bildet, abgesehen vom wenig beständigen *Ozon* O_3, im Einklang mit der *Doppelbindungsregel* zweiatomige Moleküle und damit ein Gitter entsprechend dem der Halogene. Die höheren Chalkogene bilden in kondensierter Phase keine Doppelbindungen aus, sondern Ketten oder Ringe.
Die bei Raumtemperatur stabile Modifikation des *Schwefels* ist die rhombische mit S_8-Ringen, dagegen bilden *Selen* und *Tellur* gewinkelte Ketten mit dreizähliger Schraubenachse. Die durch jeweils drei Chalkogenatome eingeschlossenen Winkel zeigen deutlich die abnehmende Bindungshybridisierung mit wachsender Ordnungszahl der Elemente. Während Schwefel mit 108° dem für sp^3-Orbitale erwarteten Wert von 109°28′ nahe kommt, ist bei *α-Polonium* mit 90° die Bindung allein schon durch das Wirken der zueinander orthogonalen *p*-Funktionen zu verstehen. Nur bei diesem Element liegt bei Normalbedingungen ein kubisch-primitives Gitter vom Bravaistyp *P* (Tafeln 70/71) vor; der Raumerfüllungsgrad ist hier nur 52 %.
V. Hauptgruppe: Mit Ausnahme von *Stickstoff*, der wegen seiner Befähigung zur Ausbildung von $p\pi$—$p\pi$-Bindungen ein Gitter aus N_2-Molekülen bildet, sind die Atome der übrigen Elemente im Gitter jeweils von drei nächsten Nachbarn umgeben. Im bei Normalbedingungen thermodynamisch stabilen schwarzen *Phosphor* liegen gewellte Doppelschichten, also ein recht hoher Vernetzungsgrad der Phosphoratome vor. Im weißen Phosphor dagegen besetzen jeweils vier Atome die Eckpunkte regulärer Tetraeder, die ihrerseits zu einem van-der-Waals-Gitter zusammentreten; die Schmelztemperatur liegt hier mit 44 °C sehr niedrig. Bei den Schichtengittern von *Arsen*, *Antimon* und *Bismut* ergibt sich eine zunehmende relative Annäherung drei weiterer Atome an das Zentralatom und seine drei nächsten Nachbarn, das *P*-Gitter mit der Koordinationszahl 6 wird aber erst bei den Hochdruckmodifikationen von Phosphor und Antimon erreicht.
IV. Hauptgruppe: *Kohlenstoff* tritt vor allem als *Graphit* und als *Diamant* auf. Das *Diamantgitter* (Tafeln 70/71) mit seinen über alle Ecken verknüpften C_4-Tetraedern läßt sich auch auffassen als kubisch flächenzentriertes Gitter, bei dem alternierend jeder zweite der acht Würfel, in die die Elementarzelle aufgeteilt werden kann, durch ein weiteres C-Atom zentriert wird. Die hohe thermische Stabilität von Diamant (↑ 6.4.3.3.) und seine extrem große Härte stehen anscheinend in Widerspruch zur geringen Raumerfüllung von 34 %, die sich für die dem Diamantgitter entsprechende Kugelpackung ergibt. Bei einer so niedrigen Koordinationszahl wie 4 kann aber das Konzept sich gerade berührender Kugeln nicht aufrechterhalten werden. Da sich die sp^3-Zustände benachbarter Atome beträchtlich überlappen, müßte man Kugeln mit einem

viel höheren „Wirkungsradius“ annehmen, die einander beträchtlich durchdringen und somit eine erheblich bessere Raumerfüllung ergäben (s. a. 6.4.4.2.).
Während *Silicium*, *Germanium* und das graue *α-Zinn* ebenfalls den Diamanttyp ausbilden, liegt bei dem weißen metallischen *β-Zinn* ein tetragonal verzerrtes Gitter vor, für das angenähert die Koordinationszahl 6 (= 4 + 2) gilt. *Blei* hat die 12er Koordination typischer Metalle.
Den Metallgittern mit ihrer sehr weitgehenden Elektronendelokalisierung stehen solche Atomgitter nahe, in denen ein Teil der Elektronen nicht fest an bestimmte Atome gebunden ist. Im Graphit (Tafeln 70/71) können die nicht an der sp^2-Hybridisierung beteiligten p-Elektronen, eins je C-Atom, über das σ-Bindungsgerüst so miteinander wechselwirken, daß innerhalb der Graphitschichten metallische Leitfähigkeit auftritt. Da zwischen den Schichten im wesentlichen van-der-Waals-Kräfte wirken, ergibt sich eine starke Anisotropie der Eigenschaften, z. B. der elektrischen Leitfähigkeit oder der Härte.

6.4.4.2. Größe und Ladung von Kristallbausteinen. Seit der Entwicklung der Röntgenstrukturanalyse (↑ 6.4.2.) sind die experimentell bestimmten Abstände r_0 zwischen den Atomkernen Grundlage für die Festlegung von *Atom-* bzw. *Ionenradien*. Bei der Aufteilung von r_0 in zwei Beiträge, die den Platzbedarf der an der Bindung r_0 beteiligten Teilchen angeben sollen, ist man aber grundsätzlich auf Hypothesen angewiesen. Meist wird vorausgesetzt, daß im Gitter sphärische Atome bzw. Ionen vorliegen, die sich gerade berühren.
Die Radien für Metalle und kovalent gebundene

Tab. 6.4.4-2 Teilchenradien bzw. partielle Bindungslängen für einige ausgewählte Hauptgruppenelemente; alle Angaben in Picometer, *KZ* Koordinationszahl

Element	Ionenradius und -wertigkeit (Goldschmidt u. a.)		metallischer Radius; *KZ* 6–12 (Pauling u. a.)	kovalenter Radius, *KZ* 4	van-der-Waals- (Wirkungs-) radius	Orbitalradius (Waber/Cromer u. a.)		ionisch-kovalenter Radius (Slater/Lebedev)
H	–	1– 135–210	–	30; 80	120	1s 53		–
Li	1+ 60–80		150–155	135	–	2s 159	1s 19 M⁺	155
Na	1+ 95–100		185–190	155	–	3s 171	2p 28 M⁺	185
K	1+ 133		255–235	195	–	4s 216	3p 59 M⁺	225
Mg	2+ 65–80		155–160	140	–	3s 128	2p 25 M⁺⁺	160
Ca	2+ 95–105		190–195	–	–	4s 169	3p 54 M⁺⁺	190
Sr	2+ 110–125		210–215	–	–	5s 184	4p 68 M⁺⁺	210
Ba	2+ 130–145		220–225	–	–	6s 206	5p 87 M⁺⁺	220
B	3+ 20–25		–	80–90	–	2p 78		–
Al	3+ 45–55		140–145	125–130	–	3p 131		120
C	4+ 15–20	4– 260	–	75–85	125–135	2p 60		77
Si	4+ 40–45	4– 200–270	135	117	170–200	3p 107		–
Sn	4+ 65–75	2+ 95–100	160	140	–	5p 124		140–155
Pb	4+ 70–85	2+ 120–130	160–175	145	–	6p 122		175–210
N	3+ 15	3– 150–170	–	70–80	150	2p 49		75
P	3+ 45	3– 185–210	–	110	190	3p 92		110
As	3+ 60–70	3– 190–220	135–150	115–120	200	4p 100		125
O	–	2– 130–145	–	65	130–140	2p 41		50
S	4+ 35	2– 175–190	–	105	155–185	3p 81		105
Se	4+ 70	2– 190–200	160	115	200	4p 92		115
Te	4+ 70–90	2– 210–220	170	130–135	220	5p 121		135
F	–	1– 130–135	–	65	135	2p 36	37 X⁻	43
Cl	5+ 35	1– 180	–	100	180	3p 73	74 X⁻	95–100
Br	5+ 45	1– 185	–	115	195	4p 85	86 X⁻	115
I	–	1– 215–220	–	135	215	5p 104	106 X⁻	135

Elemente (Tab. 6.4.4-2) erhält man durch Halbieren der Abstände r_0 (Abb. 6.4.4-3a, c); bei Atombindungen ist dabei auf deren Einfachbindungscharakter zu achten.
Bei Verbindungen teilt man traditionell die Abstände *A–B*, *B–C* usw. unterschiedlich auf, je nachdem, ob man den Stoff als ionisch oder als kovalent gebaut anspricht. Bei Ionenkristallen sieht man die Lage des Elektronendichteminimums längs der Verbindungslinie r_0 als ungefähre ‚Grenze' zwischen den unterschiedlich geladenen Teilchen an (Abb. 6.4.4-3b). Es ergeben sich relativ kleine Radien für Kationen, ziemlich große für Anionen, wie der Vergleich der isoelektronischen Teilchen Li^+/H^-, Na^+/F^-, K^+/Cl^-, Mg^{2+}/O^{2-} zeigt (↑ Tab. 6.4.4-2). Die Radienverminderung bei der Bildung edelgasähnlicher Kationen aus den Atomen erscheint wegen der Entleerung der äußeren Schale noch plausibel; verwundern muß heute aber die Annahme einer entsprechenden Radienvergrößerung bei der Bildung edelgasähnlicher Anionen, wie sie aus indirekten Messungen durch Landé und Wasastjerna (1920/23) abgeleitet wurde. Beim Zutritt von ein oder zwei *p*-Elektronen zu den im Fluor- bzw. Sauerstoffatom schon vorhandenen wird schließlich keine neue Schale besetzt.
Bei kovalent gebundenen Festkörpern ist die Grenze zwischen den Atomen ein Gebiet beträchtlicher Überlappung der bindenden Elektronenzustände (Abb. 6.4.4-3c). Das Wort Atomradius wird den Tatsachen noch weniger gerecht als der Begriff Ionenradius. Zur Beschreibung von Strukturen kovalent gebundener Stoffe verwendet man deshalb besonders in der organischen Chemie die *Wirkungs-* oder *van-der-Waals-Radien* (↑ Tab. 6.4.4-2), die zu einander durchdringenden Atomkugeln führen. Darauf bauen die *Kalottenmodelle* nach Stuart und Briegleb auf, mit denen z. B. Molekülkonfiguration und -konformation veranschaulicht werden kann (s. a. 5.4.5., 5.4.6.).
Durch die zueinander inversen Änderungen bei Kat- bzw. Anionisierung bleibt die Summe zweier konventioneller „kovalenter" Radien zugleich auch etwa die Summe zweier Ionenradien derselben Elemente (vgl. z. B.: Al/I und Al^{3+}/I^- oder Si/O und Si^{4+}/O^{2-} in Tab. 6.4.4-2). Die Übereinstimmung berechneter Radiensummen mit gemessenen Abständen r_0 kann also nicht als Kriterium für das Vorliegen des einen oder anderen Bindungstyps verwendet werden.
Durch die Befunde der Wellenmechanik (↑ 2.3.) wurde die Angabe exakter Zahlenwerte für die Radien an sich gegenstandslos. In neueren Konzepten wird als Teilchenradius der Abstand vom Kern vorgeschlagen, für den die Aufenthaltswahrscheinlichkeit der betreffenden Valenzelektronen im Grundzustand maximal ist (Waber/Cromer 1965). Auch für angeregte Zustände verwendbar sind die ähnlichen, jedoch empirisch ermittelten Werte von Slater (1964) und Lebedev (1966/69). Diese weitgehend einheitlichen Atom-Ionen-Radien (↑ Tab. 6.4.4-2) vermeiden viele Ungereimtheiten klassischer Konzepte, verlangen aber ein so krasses Umdenken bei vielen Vorstellungen, daß sie sich bis heute nicht durchgesetzt haben. Für die Raumerfüllung von SiO_2 durch Silicium und Sauerstoff sagen z. B. die Lebedevwerte aus, daß

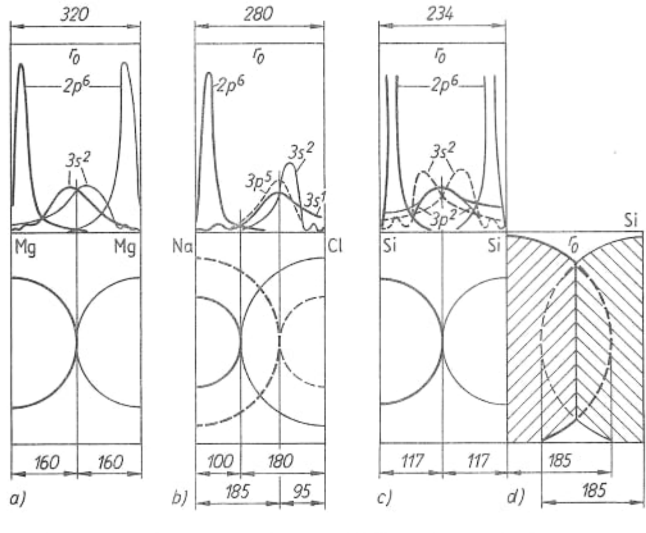


Abb. 6.4.4-3 Ausgewählte Radiensysteme und ihr Zusammenhang mit der Radialverteilung äußerer Elektronen der Bindungspartner über den Abstand r_0;
a) Magnesium („metallischer" Radius),
b) Natriumchlorid: Ionenradien nach Goldschmidt *durchgezogen*, Radien nach Lebedev *gestrichelt*,
c) Silicium („kovalenter" Radius), d) Silicium (van-der-Waals-Radius); Werte in Picometer

etwa 14% dem Sauerstoff und 86% dem Silicium zuzuschreiben seien. Dagegen ordnen die klassischen Ionenradien fast 99% des von den Teilchen eingenommenen Raums dem Sauerstoff und nur gut 1% dem Silicium zu.

Ein weiteres Beispiel: *Lithiumfluorid* kristallisiert im NaCl-Typ mit einer Gitterkonstante $a_0 = 401{,}8$ pm. Der Li—F-Abstand von $r_0 = 201$ pm stimmt recht gut überein mit der Summe der Lebedevradien, aber auch mit der Summe der klassischen „kovalenten" bzw. Ionenradien. Der wesentliche Unterschied besteht in der „Aufweichung" des Modells von den starren, kugeligen Teilchen. In der Bindung Li—F, die vom 2*s*-Zustand des Lithiums und vom 2*p*-Zustand des Fluors ausgeht, ist der Ladungsübergang des 2*s*-Elektrons zum Fluor weitgehend vollzogen (Tab. 6.4.4-4). Bei völliger Übernahme des 2*s*-Elektrons durch Fluor wäre also bei Lithium erst in sehr viel größerer Kernnähe eine relativ hohe Elektronendichte zu verzeichnen, nämlich bei etwa 19 pm, entsprechend dem für Li^+ geltenden 1*s*-Orbitalradius. Dem steht nur eine äußerst geringe Radienaufweitung beim Übergang $F \rightarrow F^-$ gegenüber (↑ Tab. 6.4.4-2). Es ergibt sich statt der dichten Packung von Teilchen also eine relativ offene Kristallstruktur.

Ein Vorteil der neuen Radien ist trotz der auch bei ihnen gebotenen Vorsicht z. B. darin zu sehen, daß sich für das Verständnis der Festkörperreaktivität, der Diffusion und der Hydratation geringere Denkschwierigkeiten ergeben, als bei der Vorstellung von sich gegenseitig berührenden Ionen. Das gilt vor allem dann, wenn man außerdem statt der meist unrealistischen Ionenladungen z (↑ 6.4.3.) die experimentell gesicherten *effektiven Ladungen* z_{eff} in Rechnung stellt.

Die direktesten Aussagen zu z_{eff} liefern die chemischen Verschiebungen in *Röntgenspektren* (↑ 4.3.3.2.3.) sowie die Bestimmung der Elektronendichteverteilung durch Röntgenkristallstrukturanalyse. Außerdem lassen sich entsprechende Korrelationen mit thermochemischen Daten, Dipolmomenten, magnetischen Suszeptibilitäten, dielektrischen, piezoelektrischen, elastischen und weiteren Konstanten aufstellen. Unterschiedliche Methoden zur Ermittlung der effektiven Ladung ergeben auf 0,1 bis 0,2 Elementarladungen übereinstimmende Werte. Für *Zinksulfid* wurde z. B. nach fünf Methoden $z_{eff(Zn)} = +0{,}67 \pm 0{,}15$ bestimmt, aus drei Messungen ein $z_{eff(S)} = -0{,}69 \pm 0{,}11$. Die durch die Oxydationszahl gegebenen Grenzwerte +2 bzw. −2 werden also bei weitem nicht erreicht. Daß eine solche Aussage auch noch bei weitaus polareren Substanzen gilt, wird aus Tab. 6.4.4-4 deutlich.

Trotz der nachgewiesenen Unzulänglichkeiten älterer Konzepte steht außer Zweifel, daß sich die Annahme konstanter Teilchenradien und -ladungen für eine erste Einführung in die geometrische Kristallchemie und für die Stoffsystematisierung durchaus bewährt hat und daraus abgeleitete Kriterien wie die *Radienquotientenregel* (↑ 6.4.4.3.) für die Lehre noch einigen Bestand haben werden.

Tab. 6.4.4-4 Effektive Ladungen z_{eff} von Hauptgruppenelementen in Vergleich mit deren Oxydationszahlen in ausgewählten Verbindungen

Verbindung	Teilchen	z_{eff}	Verbindung	Teilchen	z_{eff}
NaH	$\overset{-1}{H}$	−0,65	$NaClO_4$	$\overset{+7}{Cl}$	+2,1
LiH		−0,6	$NaClO_3$	$\overset{+5}{Cl}$	+1,5
LiF	$\overset{-1}{F}$	−0,75	$NaClO_2$	$\overset{+3}{Cl}$	+0,5
AlF_3		−0,45	NaClO	$\overset{+1}{Cl}$	±0
SiF_4		−0,4	NaCl	$\overset{-1}{Cl}$	−0,8
KCl	$\overset{-1}{Cl}$	−0,85	Na_2S	$\overset{-2}{S}$	−0,7
$CaCl_2$		−0,7	Na_2SO_3	$\overset{+4}{S}$	+1,2
$AlCl_3$		−0,45	Na_2SO_4	$\overset{+6}{S}$	+2,5
Li_2O	$\overset{-2}{O}$	−1,3	SF_6	$\overset{+6}{S}$	+3,5
BaO		−1,05	P_2S_5	$\overset{+5}{P}$	+0,9
SiO_2		−0,8	Mg_3P_2	$\overset{-3}{P}$	−0,7
SiC	$\overset{-4}{C}$	−0,7			

6.4.4.3. Strukturtypen bei chemischen Verbindungen. Da hier mindestens zwei verschiedene Atomsorten zusammentreten, ergeben sich bei deutlichen Unterschieden in Größe und Elektronegativität der Teilchen Besonderheiten gegenüber den Elementstrukturen.

Kaum verändert ist die Situation bei van-der-Waals-Gittern wenig polarer organischer und anorganischer Verbindungen. Hier wird die Gittersymmetrie maßgeblich durch die Molekülform bestimmt; meist resultieren niedrig symmetrische Strukturen.

Eine ebenfalls meist nur geringfügige Modifizierung der Elementgitter tritt bei Mischkristallbildung auf, z. B. bei Einführung von Atomen eines zweiten Elements *B* in Metallgitter *A*. Sind die Atome *B* in ihrem Raumbedarf den Atomen *A* ähnlich, können sie deren Gitterplätze einnehmen; es entstehen *Substitutionsmischkristalle* (s. a. 6.4.5.3.3.). Bei starken Größenunterschieden kommt es zu *Einlagerungsmischkristallen*, in denen die kleineren Teilchen, etwa Wasserstoff, Kohlenstoff und Stickstoff, Zwischengitterplätze der Metallstruktur einnehmen. *Palladium* kann z. B. bei Raumtemperatur das 800fache seines Eigenvolumens

an Wasserstoff aufnehmen; das Produkt hat dann etwa die Zusammensetzung „Pd_2H“.
Während Mischkristalle keine definierten chemischen Verbindungen, sondern feste Lösungen der einen Komponente in der anderen sind, kommt es bei hinreichenden Größen- und/oder Elektronegativitätsunterschieden auch bei Metallen zur Verbindungsbildung. *Cementit* Fe_3C, ein Carbid des Eisens, ist hierfür ein Beispiel (s. a. 6.4.5.).
Die große Mannigfaltigkeit der *intermetallischen Verbindungen* läßt sich nur schwer klassifizieren. Wichtige Beispiele sind: Lavesphasen mit der Bruttoformel AB_2, wobei das optimale Radienverhältnis r_A/r_B den Wert 1,22 annimmt. Die Atome A bauen ein Gitter vom Diamanttyp auf, in dessen freie Achtelwürfel die Atome B als eckenverknüpfte B_4-Tetraeder zwischengelagert sind. Substanzbeispiele sind $MgCu_2$, $MgZn_2$ und $MgNi_2$.
Zintlphasen zeigen schon einen nennenswerten Ionencharakter, so daß in einigen Fällen der Nachweis von Kationen und komplexen Anionen geglückt ist, etwa bei Na_4Sn_9, Na_3Bi_5, Na_3Sb_7. Die Struktur der Zintlphase NaTl läßt sich bei Annahme einer Grenzstruktur Na^+Tl^- einfach verstehen. Das Ion Tl^- mit seinen vier Valenzelektronen baut wie Kohlenstoff im Diamantgitter ein Tetraedernetz auf, in dessen Lücken die Na^+-Ionen Platz finden.
Hume-Rothery-Phasen lassen sich nach dem Quotienten aus der Zahl der an das Gitter abgegebenen Valenzelektronen und der Gesamtzahl der Atome einteilen. Für den Wert 21 : 14 ergeben sich kubisch raumzentrierte Phasen (Cu^IZn^{II}, $Cu^I_5Sn^{IV}$, Fe^0Al^{III}), für 21 : 13 eine besondere kubische Struktur ($Cu^I_5Zn^{II}_8$, $Pt^0_5Zn^{II}_{21}$) und für 21 : 12 eine hexagonal dichteste Packung ($Ag^I_5Al^{III}_3$, $Cu^I_{13}Sb^V_3$); die durch die römischen Ziffern angezeigte Zahl der an das Gitter abgegebenen Valenzelektronen steht in gesetzmäßiger Beziehung zur Stellung der Elemente im Periodensystem. Eine Interpretation der besonderen Stöchiometrien der Hume-Rothery-Phasen ist auf der Basis von *Bändermodell* und *Brillouinzonentheorie* möglich (s. a. 8.2.).
In Ionengittern treten wie bei Metallen Wechselwirkungen zwischen positiv und negativ geladenen Teilchen auf, jedoch sind auch die Anionen auf bestimmte Gitterplätze lokalisiert. Für die Beschreibung der wichtigsten Ionengitter empfiehlt es sich, von der kubisch bzw. hexagonal dichtesten Kugelpackung auszu-

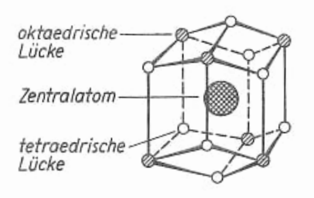


Abb. 6.4.4-5 Verteilung oktaedrischer und tetraedrischer Lücken um ein Zentralatom in einer kubisch dichtesten Packung

Tab. 6.4.4-6 Radienquotientenregel für Verbindungen des Formeltyps AB; K Kation, A Anion; *kd*, *hd* kubisch- bzw. hexagonal dicht

r_K/r_A bzw. r_A/r_K	Koordinations- -zahl	-figur	Strukturtyp	Beispiele (erwartete Koordination in Klammern)
0 → 0,153	(2)	(Hantel)	–	–
0,153 → 0,225	3	Dreieck	BN (*Bornitrid hd*)	(3) BN
0,225 → 0,415	4	Tetraeder	ZnS (*Zinkblende kd*, *Wurtzit hd*)	(3) BeS, BeTe (4) MgTe, BeO (6) ZnS, AgI, CdS
0,415 → 0,732	6	Oktaeder	NaCl (*Steinsalz kd*)	(4) LiI, LiBr, ZnSe (6) LiF, LiCl, NaF, NaCl, NaBr, NaI, KBr, KI, MgO, CaS, SrS (8) KF, KCl, RbF, RbCl, CsF, AgF, CaO, SrO, BaO
0,732 → 1	8	Hexaeder (Würfel)	CsCl (*Caesiumchlorid*)	(6) NH_4I, TlI (8) CsCl, CsBr, CsI
1	(12)	(*kd*, *hd*)	–	–

gehen und die unterschiedliche Besetzung der vorhandenen Lücken je nach Teilchengröße und Stöchiometrie der Verbindung zu betrachten. Obwohl die entstehenden Strukturen oft weit von dichtesten Packungen entfernt sind, ist so eine einfache und einprägsame Klassifizierung möglich.

Verbindet man die gedachten tetraedrischen und oktaedrischen Schwerpunkte der Lücken, die sich um ein beliebig herausgegriffenes Teilchen in einer kubisch dichtesten Packung ausbilden, erhält man ein *Rhombendodekaeder* (Abb. 6.4.4-5). Die acht Tetraederlücken entfallen zu je 1/4 auf das Zentralatom, die sechs Oktaederlücken zu je 1/6, so daß insgesamt je Atom zwei Tetraederlücken und eine Oktaederlücke verfügbar sind, ohne daß alle zugleich besetzt werden könnten.

Nach der sogenannten *Radienquotientenregel* wird durch das Größenverhältnis zwischen den Teilchen A und B festgelegt, welche Lücken besetzt, d. h., welche Koordinationspolyeder ausgebildet werden können. Unter der Annahme der Berührung aller benachbarten Ionen untereinander ergeben sich Grenzwerte der Radienquotienten r_K/r_A bzw. r_A/r_K für die Bevorzugung einer bestimmten Koordinationszahl (Tab. 6.4.4-6). Der Wert kann – unter Aufweitung der dichtesten Packung der größeren Teilchen – überschritten werden, bis eine höhere Koordinationszahl energetisch begünstigt ist. Bei Unterschreiten des Grenzwerts wäre der Ort des kleineren Ions in der Lücke nicht fixiert, eine Stabilisierung zugunsten der kleineren Koordinationszahl wäre die Folge. Die Koordinationszahlen 2 und 12 können in Ionenkristallen nicht realisiert werden.

In der Praxis wird die Radienquotientenregel etwa ebenso oft befolgt wie durchbrochen. Tab. 6.4.4-6 spricht für eine gewisse Bevorzugung der Sechserkoordination, andererseits sind auch Beispiele angegeben, wo die erwartete NaCl-Struktur nicht auftritt. Moderne Interpretationen versuchen, die auftretenden Koordinationspolyeder in Kristallen ebenso wie bei Molekülverbindungen über die Symmetrie der Bindungsorbitale zu erklären. Tetraedrische Koordination spräche dann für sp^3-Hybridisierung der Valenzelektronen der Gitterbausteine, die Koordinationszahl 6 für eine Dominanz von p- oder Mitbeteiligung von d-Zuständen; Achterkoordination wäre dort begünstigt, wo die Londonwechselwirkung zwischen den Teilchen erheblich wird (s. a. 6.4.3.1.).

Tab. 6.4.4-7 führt wesentliche Strukturtypen und ihren genetischen Zusammenhang mit den dichtesten Kugelpackungen auf. Die in Tafel 70/71 unter anderem dargestellte *Calcitstruktur* enthält planare Carbonationen, durch deren parallele Ausrichtung zueinander eine ausgeprägte optische Anisotropie bewirkt wird. Nachfolgend werden weitere wichtige Strukturtypen beschrieben.

Beim *Caesiumchlorid* bilden Chloridionen ein kubisch primitives Gitter, Caesiumionen besetzen jeweils die 8er Lücke im Würfelzentrum; man kann also von zwei ineinandergeschobenen P-Gittern sprechen. Der $Cs^{[8]}Cl^{[8]}$-Typ ist z. B. bei CsCl, CsBr, CsI, RbBr, RbI, NH_4Br und CuZn verwirklicht.

Rutil TiO_2 steht in seiner Struktur den dichtesten Kugelpackungen mit hälftiger Oktaederbesetzung nahe, kristallisiert aber im tetragonalen Kristallsystem. Weitere Vertreter des $Ti^{[6]}O_2^{[3]}$-Rutiltyps sind MgF_2, SnO_2, PbO_2, GeO_2 und das monoklin verzerrte VO_2.

Siliciumdioxid SiO_2 kommt in mehreren Modifikationen vor, von denen *Quarz*, *Cristobalit* und *Tridymit* die geläufigsten sind. In allen diesen Fällen liegen im Gitter eckenverknüpfte SiO_4-Tetraeder vor, so daß der Typ $Si^{[4]}O_2^{[2]}$ resultiert. Das Cristobalitgitter läßt sich von der Diamantstruktur herleiten. Die Si-Atome sitzen auf den Kohlenstoffplätzen, und in jede Si—Si-Bindung wird ein Sauerstoffatom eingeschoben. Die Tridymitstruktur stimmt mit der in den Tafeln 70/71 für *Eis* angegebenen Anordnung überein. Die OH_4-Tetraeder, gebildet durch je zwei kovalente O—H-Bindungen und zwei Wasserstoffbrückenbindungen je H_2O-Molekül, treten zu einer relativ offenen Struktur zusammen. Ihr weitgehendes Aufbrechen beim Schmelzen des Eises macht die Tatsache verständlich, daß sich Wasser im Gegensatz zu den meisten anderen Substanzen beim Erstarren ausdehnt.

SiO_2 ist als Endprodukt der Kondensation der *Orthokieselsäure* $Si(OH)_4$ aufzufassen. Ein Orthosilicat, der *Olivin* $MgFe(SiO_4)$, wurde schon in Tab. 6.4.4-7 erwähnt. Zwischen ihm und SiO_2 existieren eine Fülle von Silicaten mit endlichen Baugruppen, also Ketten oder Ringen, bzw. mit unendlich ausgedehnten Ketten-, Band- oder Schicht- (Blatt-) Strukturen. Obwohl sie alle durch einen analogen Vorgang, die Eckenverknüpfung von SiX_4-Tetraedern (X = OH, O) im Zuge der Kondensationsreaktion

$$\geq Si—OH + HO—Si \leq \;\longrightarrow\; \geq Si—O—Si \leq + H_2O$$

entstehen, ergibt sich eine außerordentlich große Mannigfaltigkeit durch den möglichen Eintritt von zusätzlichen Anionen oder Neutralteilchen (OH^-, F^-, O^{2-}, H_2O) in die Silicatstrukturen sowie durch die isomorphe Vertretbarkeit von Silicium durch andere geeignete Elemente, vor allem durch Aluminium. Einige Konstitutionsformeln sollen die Bauprinzipien einfacher, von zusätzlichen Anionen freier Silicate verdeutlichen: *Thortveitit* $Sc[Si_2O_7]$, *Benitoit*

Tab. 6.4.4-7 Verbindungsstrukturen, die sich von dichtesten Kugelpackungen mindestens einer Teilchensorte herleiten; *kd* kubisch dicht, *Sch* Schichtengitter; Koordinationszahlen in eckigen Klammern

Formeltyp	*B* kubisch dicht gepackt Besetzung Oktaeder-Lücken	Tetraeder-Lücken	Strukturtyp	Beispiele	*B* hexagonal dicht gepackt Besetzung Oktaeder-Lücken	Tetraeder-Lücken	Strukturtyp	Beispiele
AB	–	1/2	*Zinkblende* $Zn^{[4]}S^{[4]}$	CuF, CuCl, ZnSe, HgTe, CdS, GaP, AgI	–	1/2	*Wurtzit* $Zn^{[4]}S^{[4]}$	ZnO, CdS, AgI CrO, AlN, NH_4F
	1/1	–	*Steinsalz* $Na^{[6]}Cl^{[6]}$	LiF, BaO, KBr, NaH, TiO, VC, CrN, PbS, LaBi, RbAu	1/1	–	*Nickelarsenid* $Ni^{[6]}As^{[6]}$	TiS, $Fe_{0,90-0,93}S$, $Ni_{0,51-1,0}Te$, AuSn, $Ni_{0,85-1,18}Sb$
AB_2	1/2	–	*Cadmiumchlorid (Sch)* $Cd^{[6]}Cl_2^{[3]}$	$MgCl_2$, $FeCl_2$, $MnCl_2$, β-$TaSe_2$, Cs_2O, NiI_2	1/2	–	*Brucit (Sch)* $Mg^{[6]}(OH)_2^{[3]}$	CdI_2, PbI_2, $TiSe_2$, $Ca(OH)_2$ $Co(OH)_2$, $Mn(OH)_2$, $TiCl_2$, PtS_2
	–	1/1	*Fluorit* $Ca^{[8]}F_2^{[4]}$	SrF_2, CdF_2, PbF_2; anti-Fluorit: Li_2O, Cu_2S				
A_2B_3					2/3	–	*Korund* α-$Al_2^{[6]}O_3$	Fe_2O_3 (*Hämatit*), Cr_2O_3, V_2O_3, α-Ga_2O_3, Rh_2O_3, Ti_2O_3
AMB_3	1/4 (*A*)	–	*Perowskit* $CaTi^{[6]}O_3$ (*A* = Ti, *M* + *Bkd*)	$BaTiO_3$, $LiNbO_3$, $NaWO_3$, $YAlO_3$, $KMgF_3$, $CsAuCl_3$, $CaZrS_3$; $CdTiO_3$ (höhere Temp.)	1/3 (*A*) +1/3 (*M*)	–	*Ilmenit* $Fe^{[6]}Ti^{[6]}O_3$	$MnTiO_3$, $MgTiO_3$, $CdTiO_3$ (niedr. Temp.)
AM_2B_4	1/2 (*M*)	1/8 (*A*)	*Spinell* $Mg^{[4]}Al_2^{[6]}O_4$	$FeAl_2O_4$, $TiZn_2O_4$ *Sulfospinelle* AM_2S_4	1/2 (*M*)	1/2 (*A*)	*Olivin* $M_2^{[6]}Si^{[4]}O_4$ (*M* = Mg + Fe)	Al_2BeO_4 (Chrysoberyll), Na_2BeF_4
	1/4 (*A*) +1/4 (*M*)	1/8 (*M*)	„inverser" *Spinell*	$MgFe_2O_4$, Magnetit Fe_3O_4 [= $\overset{+3}{Fe}(\overset{+2}{Fe}\overset{+3}{Fe})O_4$]	–	1/3 (*A*) +2/3 (*M*)	*Phenakit* $Be_2^{[4]}Si^{[4]}O_4$	Zn_2SiO_4 (*Willemit*), Li_2MoO_4 Li_2WO_4, Li_2BeF_4

$BaTi(Si_3O_9]$, *Beryll* $Be_3Al_2[Si_6O_{18}]$, *Diopsid* $MgCa[SiO_3]^{1}_{2}\infty$, *Sillimanit* $Al[AlSiO_5]^{1}\infty$, *Andalusit* $Al[AlSiO_5]^{2}\infty$, *Nephelin* $Na[AlSiO_4]^{3}\infty$, *Feldspäte* wie *Albit* $Na[AlSi_3O_8]^{3}\infty$, und *Anorthit* $Ca[Al_2Si_2O_8]^{3}\infty$.

Durch eine ähnliche Vielfalt ist die Kristallchemie weiterer zur Kondensation befähigter Oxosäuren und ihrer Verbindungen ausgezeichnet, wie in den letzten Jahrzehnten vor allem für die *Borate* und *Phosphate* belegt wurde. Demgegenüber verdanken die *Sulfate* ihren Typenreichtum nicht einer unterschiedlichen Verknüpfung von SO_4-Baugruppen miteinander, sondern deren Zusammenhalt mit anderen Kation-Sauerstoff-Polyedern und der Beteiligung von Hydratwasser oder anderen, meist anionischen Gruppierungen am Gitteraufbau.

6.4.5. Zustandsdiagramme unter Beteiligung fester Phasen

Über den Zustand von Stoffen und Stoffgemischen gibt die klassische Thermodynamik zwar Auskunft, jedoch reichen die verfügbaren Daten meist nicht zur Berechnung aller interessierenden Phasenbeziehungen aus. Deshalb ist die Ermittlung von Zustands- oder *Phasendiagrammen* immer noch wichtig, z. B. für Fragen der Stofftrennung und der Werkstoff- und Prozeßoptimierung.

Phasendiagramme beziehen sich grundsätzlich auf Gleichgewichtszustände. Die Geschwindigkeit der Gleichgewichtseinstellung sowie der Weg dorthin bleiben außer acht. Als notwendiges, wenn auch nicht ausreichendes Kriterium für die erfolgte Einstellung des Gleichgewichts kann das Phasengesetz von Gibbs genutzt werden.

6.4.5.1. Gibbssches Phasengesetz. In Mehrstoffsystemen können zwar im Prinzip beliebig viele Phasen auftreten, ein thermodynamisches Gleichgewicht ist aber nur für eine gewisse kleine Zahl P möglich. Den Zusammenhang zwischen P, der Zahl F der *Freiheiten* und der Zahl K der *Komponenten* für das *Phasengleichgewicht* gibt das Gibbssche Phasengesetz (1876) an: $P + F = K + 2$. Die Summe der Zahl möglicher Phasen und Freiheiten übersteigt danach die Zahl der Komponenten um zwei. *Phasen* sind wie üblich die homogenen, durch Grenzflächen voneinander geschiedenen Teile eines *heterogenen Systems*. *Homogene Systeme* sind grundsätzlich einphasig, also auch im polykristallinen Zustand bei lockerem Zusammenhalt der Kristallite. Als Freiheiten oder *Freiheitsgrade* bezeichnet man die Zahl der intensiven Zustandsvariablen wie Druck p, Temperatur T und Molenbruch x_i, die unabhängig voneinander ohne Änderung der Phasenanzahl variiert werden können.

Die Zahl K der Komponenten ist nicht notwendig mit der Zahl S der vorhandenen Stoffe identisch; sie ist auf die Zahl der chemischen Bestandteile einzuschränken, die zur Bildung aller Phasen des Gleichgewichtssystems ausreichen: $K = S - E$. Als *Einschränkungen* E zählen z. B. die zwischen den vorhandenen Stoffen möglichen, unabhängig voneinander verlaufenden chemischen Reaktionen. So wird das System $CaO—CO_2$ durch nur zwei der drei möglichen Species, $CaCO_3$, CaO, CO_2, überstrichen, da sich gemäß der Beziehung $CaCO_3 \rightleftharpoons CaO + CO_2$ das Gleichgewicht nachstellt: $K = S - E = 3 - 1 = 2$ und damit $P + F = 4$. Liegen alle drei Stoffe in drei verschiedenen Phasen vor, hat man wegen $F = 4 - 3 = 1$ nur eine Freiheit; z. B. stellt sich zur gewählten Temperatur zwangsläufig der zugehörige Zersetzungsdruck ein.

Oft verzichtet man auf den Gebrauch aller vorhandenen Freiheiten. Luft enthält mindestens zehn verschiedene Gase S, wird aber meist als *ein* Stoff betrachtet. Das Konstanthalten von $(S - 1)$ Molenbrüchen x_i führt dann zu ebensovielen Einschränkungen, so daß Luft tatsächlich als *eine* Komponente anzusprechen ist: $K = S - E = S - (S - 1) = 1$.

6.4.5.2. Einkomponenten- oder unitäre Systeme. Da die chemische Zusammensetzung konstant ist, kann die Abhängigkeit des Zustands von den verbleibenden Variablen p und T gleichzeitig in *einem* zweidimensionalen p-T-Diagramm dargestellt werden. Abb. 6.4.5-1 zeigt einen Ausschnitt aus dem *Zustandsdiagramm des Wassers*, der die Hochdruckformen von Eis nicht enthält. Die sich zu jeder Temperatur einstellenden *Sättigungsdampfdrücke* von Eis bzw. Wasser ergeben die Kurven A und B. Die *Sublimationsdruckkurve* A gibt das stabile thermodynamische Gleichgewicht zwischen Eis und Wasserdampf an, die *Dampfdruckkurve* B trennt das Zustandsfeld flüssigen Wassers von dem des Dampfs. Zum Sieden kommt es, wenn der Dampfdruck den Umgebungsdruck erreicht; tabellierte Siedetemperaturen T_V gelten meist

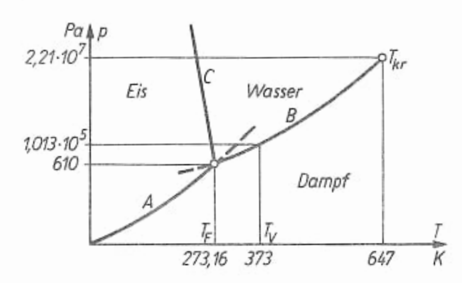


Abb. 6.4.5-1 Zustandsdiagramm des Wassers (Erläuterung im Text)

für Normaldruck von etwa 10^5 Pa (1 atm). Die Dampfdruckkurve einer Flüssigkeit endet am *kritischen Punkt* T_{kr}.
Der Schnittpunkt von A und B stellt die Schmelztemperatur T_F von Eis bzw. die Gefrier- oder Erstarrungstemperatur von Wasser unter dem eigenen Dampfdruck von 610 Pa dar. Bei dieser Temperatur von 273,16 K sind alle drei Aggregatzustände des Wassers nebeneinander beständig (*Tripelpunkt*); das Gleichgewicht ist wegen $F = 3 - P = 0$ nonvariant. Erhöht man die Temperatur, schmilzt das Eis, die Zahl der Phasen geht auf zwei zurück, F wird gemäß $F = 3 - 2 = 1$. Gibt man längs der Linie B die Temperatur T vor, nimmt p dann jeweils den durch das Gleichgewicht vorgegebenen Wert an. Läßt man Luft als Komponente zu, wird $K = 2$ und $F = 4 - P$. Die zusätzliche Freiheit besteht in der Änderung des Partialdrucks der Luft in der Gasphase. Beläßt man es beim Normaldruck, liegt de facto wieder ein Einkomponentensystem vor, wenn auch mit geringfügig veränderten Eigenschaften; z. B. erniedrigt sich die Erstarrungstemperatur T_F von luftgesättigtem Wasser auf 273,15 K, d. h. definitionsgemäß exakt auf $\vartheta_F = 0$ °C. Nach der *Schmelzdruckkurve* C nimmt T_F von Eis mit steigendem Druck nicht zu, sondern ab. Dieses ungewöhnliche Verhalten ist durch die relativ offene Eisstruktur (↑ 6.4.4.3.) bedingt.
Phasenumwandlungen erfolgen nur dann glatt, wenn Keime der neuen Phase zugeführt werden oder schon vorgebildet sind. Unterbleibt z. B. die Kristallisation bei Unterschreiten von T_F, liegt längs des gestrichelten Teils von B das thermodynamisch metastabile Gleichgewicht zwischen *unterkühlter Schmelze* und Wasserdampf vor. Beim Ausbleiben der Dampfkondensation unterhalb T_V spricht man von *übersättigtem Dampf*; der gegenläufige Fall, das Überhitzen von Wasser, wird als *Siedeverzug* bezeichnet. Wie die Unterkühlung der Schmelze durch Impfkristalle (↑ 6.4.2.1.), läßt sich der Siedeverzug durch Dampfbläschen aufheben, die sich an scharfkantigen, porösen „Siedesteinen" leichter ausbilden als an glatter Gefäßwand oder im Flüssigkeitsvolumen selbst.
Das *Sublimieren*, also die direkte Umwandlung Festkörper ⇌ Dampf, gelingt gemäß Abb. 6.4.5-1 unterhalb des Tripelpunktdrucks. Iod und Naphthalen, die bei Normaldruck sublimieren, haben folglich Tripelpunktdrücke $p_{Tr} > 10^5$ Pa. Feste Stoffe mit kleinen Werten p_{Tr} lassen sich bei entsprechend vermindertem Druck in den Gaszustand überführen oder aus ihm abscheiden.
Phasenumwandlungen im festen Zustand finden bei Eis nur unter hohem, bei vielen anderen Stoffen aber schon bei normalem Druck statt. Das Auftreten mehrerer fester Formen, der *Modifikationen*, wird bei Elementen als *Allotropie* [*allos tropos* griech., anderer Charakter], bei Verbindungen als *Polymorphie* [*polys* griech., viel, *morphe* Gestalt] bezeichnet. Modifikationswechsel führen wie Änderungen des Aggregatzustands zu Knicken in den p-T-Kurven. Umwandlungen können *enantiotrop*, d. h. wie bei Schwefel wechselweise in beiden Richtungen erfolgen, oder *monotrop*, also einseitig wie bei der Überführung von weißem in schwarzen Phosphor (Abb. 6.4.5-2).

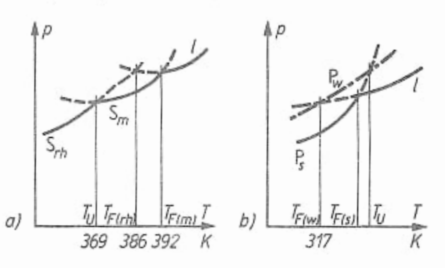


Abb. 6.4.5-2 Druck-Temperatur-Kurven (schematisch); a) bei enantiotroper (Schwefel), b) bei monotroper (Phosphor) Umwandlung; *rh* rhombisch, *m* monoklin, *s* schwarz, *w* weiß, *l* Schmelze, T_U Umwandlungs-, T_F Schmelztemperatur

Bei der Bildung eines in mehreren Modifikationen existierenden Stoffs entsteht nicht immer zuerst die energiereichste Modifikation, wie die *Stufenregel* von OSTWALD postuliert, sondern die, deren Aufbau die geringste *Keimbildungsarbeit* erfordert.
Von den zahlreichen Formen, in denen fester Schwefel existieren kann, z. B. als S_6, S_7, S_8, S_{12}, S_{18}, S_{20} und S_∞, treten im Gleichgewichtsdiagramm nur 2 in Erscheinung, eine orthorhombische S_8-Modifikation und ab 96 °C eine monokline Form S_8. Hierin äußert sich die für die Synthesechemie sehr wichtige Tatsache, daß das Auftreten einer bestimmten Phase im Diagramm wohl ein hinreichendes, aber kein notwendiges Kriterium für deren mögliche Existenz ist. In vielen Fällen sind gerade solche metastabilen Phasen für Reaktivität und andere Eigenschaften von Stoffgemischen besonders wichtig.

6.4.5.3. Zweikomponenten- oder binäre Systeme. Neben p und T muß nun auch die Zusammensetzung angegeben werden, etwa als Molenbruch x_i oder Massenbruch w_i einer Komponente. Will man bei zweidimensionalen Darstellungen bleiben, sind gesonderte x_i-p- bzw. x_i-T-Diagramme aufzustellen, bei denen die nicht erfaßte Variable, T bzw. p, jeweils konstant gehalten wird. Für Flüssig-Dampf-Gleichgewichte, die *Siedediagramme*, sind beide Darstellungen gebräuchlich, für Flüssig-Fest-Gleichgewichte, die *Schmelzdiagramme*, genügt wegen der geringen Kompressibilität der kon-

densierten Phasen und der geringen Dampfdrücke über ihnen meist das x_l-T-Diagramm.
Da bei binären Systemen A–B neben mehreren festen Phasen auch mehr als eine flüssige auftreten kann, dient meist deren Mischbarkeit im flüssigen bzw. festen Aggregatzustand zur Fallunterscheidung. Wird zunächst Verbindungsbildung zwischen A und B ausgeschlossen, sind drei Grenzfälle zu diskutieren.

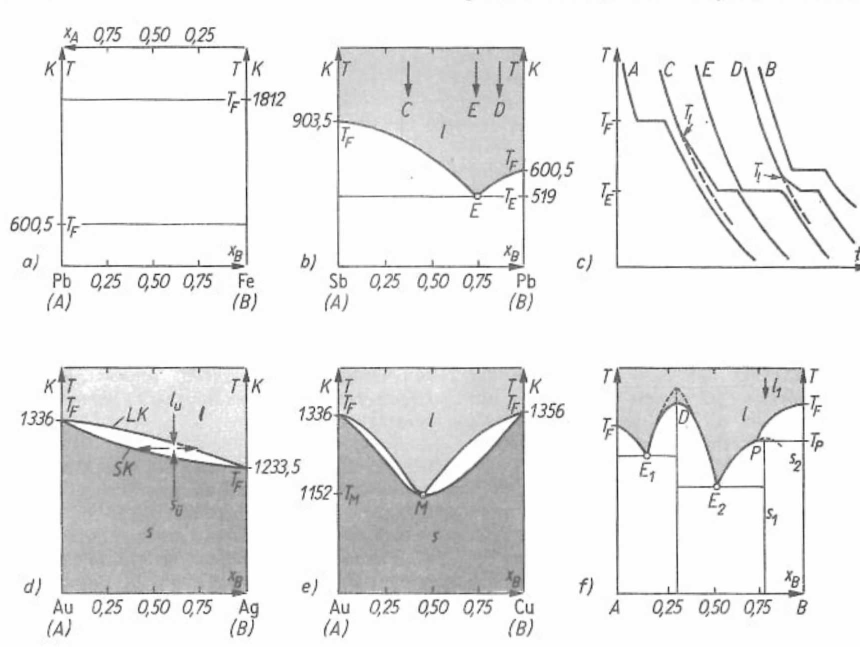


Abb. 6.4.5-3 Schmelzdiagramme binärer Systeme A–B; a) keine Mischbarkeit im flüssigen wie im festen Zustand (Blei–Eisen), b) keine Mischbarkeit im festen Zustand (Antimon–Blei), c) Abkühlungskurven für die Schmelzen A–B von b), d) ideale Mischbarkeit im flüssigen wie festen Zustand (Gold–Silber), f) Diagramm mit zwei Eutektika und der Bildung je einer kongruent (D) bzw. inkongruent (P) schmelzenden Verbindung; Zusammensetzung: Molenbrüche x_B, x_A

6.4.5.3.1. *Unmischbarkeit im flüssigen wie festen Zustand.* Da es eine absolute Unlöslichkeit einer Komponenten in einer anderen nicht gibt, steht der Begriff „unmischbar" für eine so geringe Löslichkeit, daß sie sich im Zustandsdiagramm nicht zu erkennen gibt. Die Unmischbarkeit im System Eisen – Blei (Abb. 6.4.5-3a) wird bewirkt durch die geringe chemische Verwandtschaft beider Elemente und begünstigt durch deren Dichteunterschiede. Erhitzt man ein beliebig zusammengesetztes Gemisch, so schmilzt bei $T_{F(A)}$ das Blei. Das System ist hier nonvariant, da festes und geschmolzenes Blei mit festem Eisen und Metalldampf koexistieren. Oberhalb $T_{F(A)}$ verschwindet festes Blei, das System wird monovariant, bis bei $T_{F(B)}$ mit flüssigem Eisen wieder eine vierte Phase auftritt. Oberhalb $T_{F(B)}$ liegen dann geschmolzenes Blei und Eisen neben Metalldampf vor.

6.4.5.3.2. *Unmischbarkeit im festen, Mischbarkeit im flüssigen Zustand.* Löst man in reinem A, z. B. in Antimon, etwas B, z. B. Blei, oder in reinem B etwas A, dann erniedrigt sich entsprechend dem *Raoultschen Gesetz* die Gefriertemperatur beider Schmelzen (Abb. 6.4.5-3b). Bei E, der tiefsten Temperatur für die Existenz der flüssigen Phase, schneiden sich beide Kurven. Hier scheiden sich in einer *eutektischen Reaktion*, $l \rightleftarrows s_A + s_B$, aus der Schmelze zwei feste Phasen nebeneinander unter Bildung eines *eutektischen Gemenges* aus [*eutektos* griech., gut gebaut]; dieses steht im Gleichgewicht mit Schmelze l_E und Metalldampf, so daß am *Quadrupelpunkt* E des binären Eutektikums vier Phasen koexistieren.

Aus der Schmelze C scheiden sich bei der Liquidustemperatur $T_{l(C)}$ erste Antimonkristalle aus, die sich bei weiterer Abkühlung vergrößern. Dabei reichert sich die Schmelze bis zur Zusammensetzung E an Blei an. Bei T_E werden die großen Antimonkristalle von sich nunmehr ausscheidendem feinem eutektischem Gemenge

umhüllt. Analog erhält man letztlich aus der Schmelze D große, in das Eutektikum eingebettete Bleikristalle.
In der Praxis stellt man oft Legierungen etwa eutektischer Zusammensetzung her, da dann ein „gutgebautes", feinkörniges Material resultiert. Die Koexistenz der beiden Phasen im Eutektikum läßt sich mikroskopisch und röntgenographisch belegen, z. B. durch die charakteristischen Linien der reinen Stoffe A und B in Debye-Scherrer-Aufnahmen (↑ 6.4.2.3.3.).
Zusätzliche Informationen liefert die *thermische Analyse*, in einfachen Fällen schon durch die Aufnahme der Temperatur-Zeit-Kurven für die Abkühlung der Schmelzen (Abb. 6.4.5-3c). Die Kristallisation von reinen Stoffen oder eutektischen Gemengen führt bei den jeweiligen Erstarrungstemperaturen durch die freiwerdende Kristallisationswärme zu Haltestrecken im Diagramm. Die Abkühlungskurven der Schmelzen C und D werden ab $T_{l(C)}$ bzw. $T_{l(D)}$, dem Beginn der Ausscheidung einer reinen Komponente, flacher und zeigen bei T_E ebenfalls eine Haltestrecke, die allerdings wegen der geringeren Menge verbliebener Schmelze verkürzt ist. Unterhalb T_E geben alle Kurven die Abkühlung der festen Phasen auf Umgebungstemperatur wieder.
Für Salzschmelzen mit Eutektikum ist das System *Kryolith – Aluminiumoxid*, Na_3AlF_6–Al_2O_3, ein Beispiel. Kryolith schmilzt bei etwa 1000 °C und Aluminiumoxid bei etwa 2000 °C, jedoch liegt T_E des Gemischs mit $w_{Al_2O_3} \approx 20\%$ bei 1210 K. Die Schmelzflußelektrolyse von Al_2O_3, kann also bei einer vergleichsweise niedrigen Temperatur unterhalb 1000 °C durchgeführt werden. Die *Schmelztemperaturerniedrigung* auf T_E wird auch genutzt, um Temperaturen bis zu −50 °C im Labor zu erzeugen. Eine wäßrige Kochsalzlösung mit $w_{NaCl} = 22{,}4\%$ erstarrt z. B. erst bei −21 °C. Bei 0 °C ist ein solches Eis-Kochsalz-Gemisch also nicht im Gleichgewicht; das Eis schmilzt unter Wärmeaufnahme, bis T_E erreicht ist („*Kältemischung*").

6.4.5.3.3. *Völlige Mischbarkeit im festen und flüssigen Zustand.* Hier gibt es für jeden Aggregatzustand lediglich eine Phase (Abb. 6.4.5-3d). Liegen alle drei nebeneinander vor, verbleibt eine Freiheit; durch Wahl der Zusammensetzung x_s *oder* x_l sind die übrigen Variablen festgelegt. Die unterschiedliche Zusammensetzung von Schmelze und konjugierter fester Phase kann zur Trennung der Stoffe A und B durch *fraktionierte Kristallisation* genutzt werden.
Unterhalb der *Soliduskurve* SK oder oberhalb der *Liquiduskurve* LK können T und x_l unabhängig voneinander verändert werden. Das durch beide Kurven eingeschlossene, bei idealen Mischungen linsenförmige Gebiet, verkörpert keinen stabilen Zustand. Nach Aufheben kinetischer Hemmungen zerfallen unterkühlte Schmelze l_u bzw. überhitzter Festkörper $s_ü$ in die durch x_l, x_s und T bestimmten koexistierenden Phasen. Ideale Mischbarkeit ist dann zu erwarten, wenn die Komponenten A und B einander so ähnlich sind, daß ihre Teilchen sich isomorph vertreten können, wie im Falle der Bildung von *Substitutionsmischkristallen* im System *Gold – Silber*. Bei immer noch völliger Mischbarkeit, aber schon größeren Unterschieden im chemisch-strukturellen Verhalten der Komponenten A und B treten Extrema, teils Maxima, teils Minima, in den Zustandsdiagrammen auf (Abb. 6.4.5-3e). Der Berührungspunkt M von Solidus- und Liquiduslinie stellt in Analogie zu den *azeotropen Gemischen* bei entsprechenden Siedediagrammen einen nonvarianten Zustand dar. Durch die Identität von $x_{M(l)}$ und $x_{M(s)}$ ergibt sich eine Einschränkung E und damit entsprechend der Dreiphasigkeit für $F = (K - E) + 2 - P = (2 - 1) + 2 - 3 = 0$. Die Abkühlungskurve einer solchen Mischung ist nicht zu unterscheiden von der reiner Stoffe oder eutektischer Gemenge; in Debye-Scherrer-Aufnahmen würden sich allerdings gesetzmäßige Linienverschiebungen gegenüber den Diffraktogrammen von A und B ergeben.

6.4.5.4. Phasendiagramme A–B mit Verbindungsbildung. Es sind vor allem zwei Fälle zu unterscheiden (Abb. 6.4.5-3f): Eine Verbindung, die thermisch so stabil ist, daß sie *kongruent*, d. h. ohne Änderung ihrer Bruttozusammensetzung schmilzt, ergibt im Phasendiagramm ein Schmelztemperaturmaximum D oder *Dystektikum* [*dystektos* griech., schlecht gebaut]. Dieser abwertende Name für eine doch einheitliche Verbindung rührt daher, daß intermetallische Verbindungen wegen ihrer spröden großen Kristalle als Werkstoffe oft nicht in Frage kamen. Abgeflachte Dystektika deuten auf eine partielle Dissoziation der Verbindungen beim Schmelzen hin.
Manche Verbindung s_1 ist thermisch nicht allzu belastbar. Sie schmilzt bei T_P *inkongruent*, steht also nicht mit einer Schmelze l_1 im Gleichgewicht, sondern zerfällt in einer *peritektischen Reaktion* [*peritektos* griech., um ... herumgebaut], $s_1 \rightleftarrows s_2 + l$, in neuen Festkörper s_2 und entsprechend anders zusammengesetzte Schmelze l. Wie bei einer eutektischen Reaktion stehen hier zwei feste Phasen im Gleichgewicht mit einer flüssigen. Der Name *Peritektikum* erklärt sich daraus, daß beim Abkühlen der entsprechend zusammengesetzten Schmelze l_1 zunächst Kristalle s_2 ausgeschieden werden, die sich bei T_P durch Reaktion mit der Schmelze von der Oberfläche her in Kristalle s_1 umbilden. Da nun diese Reaktion bei üblichen Abkühlgeschwindigkeiten selten zum Abschluß kommt, sind mikroskopisch im Schliffpräparat charakteristische Hülle-Kern-Gebilde zu sehen.

6.4.5.5. Eisen-Kohlenstoff-Diagramm. Nach den Grundtypen binärer Zustandsdiagramme wären noch Drei- und Vielkomponentensysteme sowie zahlreiche Varianten mit beschränkter Mischbarkeit im flüssigen oder festen Aggregatzustand zu behandeln. Einige der Besonderheiten sollen am Beispiel des *Eisen-Kohlenstoff-Diagramms* erläutert werden. Wegen der grundlegenden Bedeutung für die mehr als 10000 technisch verwendeten Eisenlegierungen ist der betreffende Bereich – von reinem Eisen bis zu wenigen Masseprozent Kohlenstoff – ausführlich untersucht worden. Das in Abb. 6.4.5-4 gezeigte Teildiagramm ist als eutektisches System mit Verbindungsbildung und beschränkter Mischkristallbildung anzusprechen. Die Verbindung Fe_3C, *Cementit*, bestimmt durch Ausbildungsform und Verteilung im Gefüge wesentlich Festigkeit, Härte und Zähigkeit des 0,05 bis 1,7 Masse-% Kohlenstoff enthaltenden *Stahls*. Die Phasenbeziehungen Fe-C komplizieren sich dadurch, daß das Gleichgewicht Fe-Fe_3C metastabil ist, also auch das technisch weniger wichtige, aber thermodynamisch stabile Gleichgewicht Eisen-Graphit zu berücksichtigen ist.

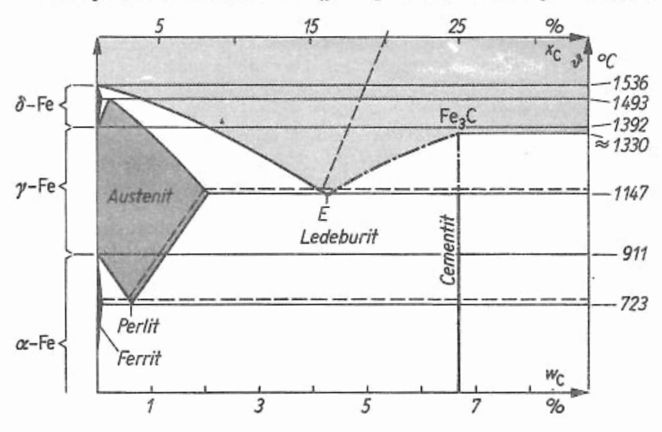


Abb. 6.4.5-4 Eisen-Kohlenstoff-Diagramm (Ausschnitt, vereinfacht); System Fe–Fe_3C *dicke Linien*, System Fe–$C_{Graphit}$ *dünne Linien*, w_C bzw. x_C Massen- bzw. Molenbruch an Kohlenstoff in Prozent
E Eutektikum

Eisen tritt in drei kubischen Modifikationen auf, dem α-Fe mit *I*-, dem γ-Fe mit *F*- und dem δ-Fe mit *I*-Gitter. Alle drei Allotrope nehmen Kohlenstoff auf, α-Fe maximal 0,02% unter Bildung von als *Ferrit* bezeichneten Mischkristallen. Geringfügig ist auch der Kohlenstoffanteil in δ-Fe, wogegen sich in γ-Fe Kohlenstoff bis zu Molenbrüchen $x_C = 8\%$ unter Gitteraufweitung einführen läßt. Die entstehende Phase wird als *Austenit* bezeichnet, ein bei rascher Abkühlung gebildetes metastabiles Umwandlungsprodukt als *Martensit*.

Stark vereinfacht ergeben sich folgende Phasenbeziehungen. Austenit kristallisiert unterhalb 1500 °C aus einer kohlenstoffhaltigen Eisenschmelze aus. Die Schmelze reichert sich an Kohlenstoff an, bis zur Sättigung bei etwa $x_C = 17\%$ und 1147 °C; es bildet sich ein zweiphasiges eutektisches Gefüge, das bis etwa 700 °C aus Cementit und Austenit besteht und als *Ledeburit* bezeichnet wird. Bei „unendlich langsamer" Abkühlung müßte sich statt Cementit Graphit ausscheiden.

Unterhalb ϑ_E sinkt die Löslichkeit von Kohlenstoff im Austenit; der beim Abkühlen freigesetzte Kohlenstoff wird unter Bildung von Sekundär-Cementit abgefangen. Bei etwa 700 °C hört Austenit auf zu existieren. In einer der eutektischen ähnlichen *eutektoiden Reaktion*, $s_1 \rightleftharpoons s_2 + s_3$, zerfällt Austenit mit 0,8 Masseprozent C in Ferrit und Cementit. Das eutektoide Gemenge aus Ferrit und Cementit trägt zwar einen eigenen Namen, *Perlit*, wie bei Ledeburit darf aber hieraus nicht auf das Vorliegen einer einheitlichen Phase geschlossen werden.

Der unter Abschreckbedingungen („Härten") aus Austenit entstehende Martensit kann als tetragonal verzerrter, kohlenstoffübersättigter Ferrit aufgefaßt werden. Durch thermische Nachbehandlung („Anlassen") des harten und spröden Martensits wird der Übergang zu Ferrit und Cementit bewirkt. Es kommt dabei zur Aufhebung innerer Spannungen, ohne nennenswerten Härteabfall, aber mit einem Gewinn an Zähigkeit. Während man Stahl auf diese Weise „vergütet", läßt sich *Gußeisen* mit 2 bis 4 Masseprozent C weder in der Wärme umformen noch vergüten. Bei Wärmebehandlung entsteht aus Primär-Cementit Graphit.

Technologisch nicht vermeidbare oder gewollte Zusätze („legierte Stähle") komplizieren das Eisen-Kohlenstoff-Diagramm und die Kinetik der Umwandlungen. Neuerdings werden *TTT-Schaubilder* [engl. *t*ime-*t*emperature-*t*ransformation] ermittelt, die das Phasengefüge in Abhängigkeit von Art und Dauer der Wärmebehandlung darstellen (s. a. 6.4.6.3.).

6.4.6. Gläser und andere nichtkristalline Stoffe

Nichtkristalline feste Körper weisen keine periodische Anordnung ihrer Baugruppen in den drei Raumrichtungen auf. Da sie folglich nicht wie Kristalle durch charakteristische Wachstumsflächen begrenzt werden, bezeichnet man sie häufig als *amorph* [*amorphos* griech., gestaltlos]. Wegen ihrer hohen Fehlordnung sind sie energiereicher, damit oft auch reaktiver als kristalline Phasen; sie besitzen eine höhere Wärmekapazität und eine endliche Nullpunktsentropie.

6.4.6.1. Einteilung nichtkristalliner Stoffe. Obwohl es zur Einteilung amorpher Stoffe keine einheitliche Auffassung gibt, ist es für viele Betrachtungen zweckmäßig, die aus der Schmelze nichtkristallin erstarrten Stoffe als *Gläser* von den übrigen amorphen Materialien abzugrenzen, die man ihrerseits grob unterteilen kann in aus dem Dampf abgeschiedene Schichten sowie in die aus kristallinen Festkörpern oder aus Lösungen erhaltenen Produkte.

Amorphe Schichten lassen sich auf vielfältige Weise herstellen, z. B. durch Abscheidung aus dem Dampf auf eine hinreichend kalte Unterlage, durch thermische Zersetzung oder Hydrolyse leichtflüchtiger Stoffe in der Dampfphase, durch Zersetzung in Glimmentladungen, auch durch „Zerstäuben“ von Festkörpern von der Oberfläche her durch das Aufprallen beschleunigter Ionen. Ein besonders großes Interesse besteht gegenwärtig an einer möglichst reproduzierbaren großflächigen Abscheidung von *amorphem Silicium*, um dadurch die Umwandlung von Sonnen- in elektrische Energie in größerem Umfang betreiben zu können, als es zur Zeit mit einkristallinem Material möglich ist. Aus Lösungen lassen sich amorphe Niederschläge ausfällen, zum Beispiel einige Sulfide von Arsen und Antimon oder Kieselsäure bzw. Kieselgel $SiO_2 \cdot x\,H_2O$; manche Stoffe lassen sich auch elektrolytisch in dünner amorpher Schicht abscheiden.

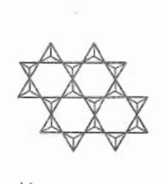
Kristalline Festkörper können durch Beschuß mit energiereichen Teilchen zumindest oberflächlich „amorphisiert“ werden, ebenso durch Feinstmahlung, Schleifen und Polieren.

Allen nichtkristallinen Stoffen ist gemeinsam, daß ihr Energiegehalt deutlich über dem der entsprechenden kristallinen Phasen liegt und daß sie trotz meist recht gut erhaltener *Nahordnung* in den jeweiligen Koordinationspolyedern keine *Fernordnung* in deren geometrischen Beziehungen zueinander mehr aufweisen (Abb. 6.4.6-1). Bei Röntgenbeugungsuntersuchungen ergeben sich statt scharfer Linien nur noch verwaschene, sich stark überlagernde Reflexe ähnlich denen, die bei der Beugung an Kristalliten sehr geringer Abmessungen auftreten.

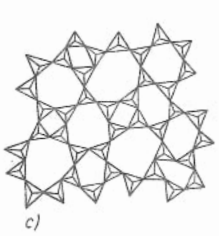
Die Unsicherheit in der Einteilung nichtkristalliner Stoffe rührt teilweise daher, daß sich unterschiedliche Herstellungsverfahren nicht auf alle Eigenschaften gleich auswirken. So ist die spezifische elektrische Leitfähigkeit kompakter Gläser oft kaum anders als die aus dem Dampf abgeschiedener amorpher Schichten, so daß hier eine Unterscheidung unnötig wäre. Andererseits sind die Oberflächenverhältnisse und die daraus abgeleiteten Eigenschaften wie Adsorption, katalytische Wirkung und Stabilität gegenüber Kristallisation oft sehr verschieden für aus dem Schmelzfluß entstandene, aufgedampfte oder im festen Zustand „amorphisierte“ Produkte.

Die durch ihre Herstellung voneinander abgegrenzten Gruppen amorpher Stoffe lassen sich weiter untergliedern nach ihrer stofflichen Zusammensetzung. Das führt bei Gläsern z. B. zu *Silicatgläsern* als dem Glas im engsten Sinne, *organischen Gläsern*, *Phosphat-*, *Chalkogenid-*, *Halogenid-* und *metallischen Gläsern*. Die Mannigfaltigkeit nichtkristalliner Stoffe, auch Ruß, Bernstein, Siegellack wären hier einzuordnen, kann nur durch die Auswahl der Beispiele angedeutet werden.

a)



b)



c)

Abb. 6.4.6-1 a) im Anion SiO_4^{4-} ist Silicium tetraedrisch von vier Sauerstoffatomen umgeben (Nahordnung), b), c) Schema der Tetraederverknüpfung bei kristallinem (Fernordnung) bzw. glasigem SiO_2

6.4.6.2. Abgrenzung der Gläser von kristallinen Stoffen und Schmelzen. Die verschiedenen Modelle, die den *Glaszustand* strukturell, thermodynamisch oder kinetisch beschreiben, durchdringen sich so sehr, daß ihre getrennte Behandlung nicht sinnvoll erscheint. Schon die auf das Zustandsdiagramm bezogene Definition „Gläser sind *eingefrorene unterkühlte Schmelzen*“, mehr noch der Satz „Glasbildung ist verhinderte Kristallisation“, implizieren, daß über die Abkühlgeschwindigkeit und die Gleichgewichtseinstellzeiten das Einbeziehen kinetsicher Vor-

stellungen unumgänglich wird. Strukturchemische Bezüge ergeben sich dadurch, daß je nach den wirkenden Bindungskräften und der Valenzelektronenkonfiguration der beteiligten Elemente unterschiedliche Bedingungen für das Unterkühlen und schließlich Einfrieren einer Schmelze gegeben sind.

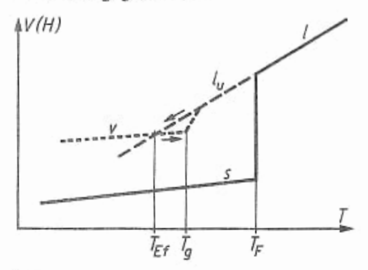


Abb. 6.4.6-2 Volumen-Temperatur- bzw. Enthalpie-Temperatur-Diagramm(schematisch) für kristallines (s) und glasartiges (v) Erstarren von Schmelzen (l); l_u unterkühlte Schmelze, T_{Ef} Einfriertemperatur, T_g Transformationstemperatur, T_F Schmelztemperatur

Abb. 6.4.6-2 zeigt in einem *Volumen-Temperatur-Diagramm* den Phasenübergang flüssig – fest. Mit abnehmender Temperatur nimmt die Schmelze l ein immer kleineres Volumen ein, bis sie bei T_F in die kristalline Phase s übergeht. Ist die Hemmung für den Phasenübergang $l \rightarrow s$ genügend groß, z. B. bei rascher Abkühlung der von Verunreinigungen freien Schmelze, kann die Kristallisation ausbleiben, die nunmehr *unterkühlte Schmelze* l_u folgt dem extrapolierten Teil der Kurve. Sie befindet sich noch in einem thermodynamischen Gleichgewicht mit der Dampfphase, allerdings in einem nur noch metastabilen; ihm kommt gegenüber geringfügigen Zustandsänderungen noch die energetisch günstigste Position zu, nicht mehr aber gegenüber der energetisch tiefer liegenden kristallinen Phase s.

Zur erneuten Gleichgewichtseinstellung nach jeder Temperaturänderung ist immer eine endliche Zeit, die *Relaxationszeit*, erforderlich, andererseits erschwert die mit fallender Temperatur zunehmende Viskosität der unterkühlten Schmelze Teilchenverrückungen immer mehr. Es wird zu einer, vom Viskositätsverlauf und der Kühlgeschwindigkeit abhängigen, *Einfriertemperatur* T_{Ef} kommen, ab der die Schmelze nicht mehr das metastabile Gleichgewicht erreicht. Erst von da an ist sie ein Glas v [*vitreus* lat., glasig] im vorgenannten Sinne. Bei der Glasherstellung und -verarbeitung wird allerdings die Bezeichnung Glas auch auf die noch gut fließ- und verarbeitungsfähigen unterkühlten Schmelzen zwischen T_{Ef} und T_F ausgedehnt.

Verfolgt man die Temperaturabhängigkeit bestimmter Glaseigenschaften von Raumtemperatur an, ergeben sich deutliche Änderungen – entsprechend dem Übergang Glas/unterkühlte Schmelze – nicht exakt bei T_{Ef}, sondern bei einer von der Heizgeschwindigkeit abhängigen höheren Glas- oder *Transformationstemperatur* T_g, oberhalb der die Volumen- oder auch Enthalpiekurve wieder in die metastabile Linie l_u einmündet. Es lassen sich also aus *einer* glasbildenden Schmelze je nach Kühl- und Heizgeschwindigkeit im Prinzip beliebig viele, mehr oder weniger deutlich voneinander unterscheidbare Gläser erhalten. Somit ist ein Glas wie auch andere nichtkristalline Stoffe durch die üblichen Zustandsvariablen der klassischen Thermodynamik, etwa Druck und Temperatur, nicht eindeutig zu beschreiben. Der im jeweiligen Glas vorhandene Ordnungszustand müßte durch zusätzliche zeitabhängige Parameter festgelegt werden. Zumindest sind genaue Angaben notwendig zu dem bei der Glasherstellung und -nachbehandlung befolgten Temperatur-Zeit-Regime.

Die Tatsache, daß Gläser vom Standpunkt der Thermodynamik in zweifacher Hinsicht instabil sind, gegenüber unterkühlter Schmelze und kristallinem Festkörper, schließt nicht aus, daß sie bei deutlich tieferen Temperaturen als T_{Ef} dennoch über große Zeiträume hinweg, bis zu Jahrtausenden, ihren Zustand im wesentlichen beibehalten, also im landläufigen Sinne durchaus stabil sind; bei den üblichen Einfrier- und Transformationstemperaturen zwischen 400 und 600 °C, bei *Kieselglas* SiO_2 sogar mehr als 1000 °C, ist das der normale Fall. Diese „Stabilität“ zeigt wiederum, daß die Thermodynamik lediglich Aussagen zur Möglichkeit von Zustandsänderungen treffen kann, wogegen der tatsächliche Ablauf durch die Reaktionskinetik bestimmt wird. Genau genommen sind auf nichtkristalline Stoffe die klassischen Zustandsfunktionen der Thermodynamik gar nicht anwendbar, da die wesentliche Voraussetzung dafür, die Einstellung des inneren Gleichgewichts, nicht gegeben ist.

6.4.6.3. Glasbildung in Abhängigkeit von kinetischen und strukturellen Faktoren. Die Kristallisation einer Schmelze unterbleibt, wenn der Temperaturbereich mit günstigen Werten der *Keimbildungsgeschwindigkeit* KB und/oder der *Kristallwachstumsgeschwindigkeit* KW genügend rasch durchlaufen wird. Abb. 6.4.6-3 zeigt schematisch die Temperaturabhängigkeit für beide Größen und für die *Viskosität* der Schmelze. Bestimmte Mindestwerte von KB und KW für das Eintreten der Kristallisation lassen sich nicht angeben. Je größer die *Unterkühlung* $\Delta T = T_F - T$, desto kleinere Kristallkeime werden für das weitere Wachstum und den Übergang zu Kristallen lebensfähig; diese

Schwelle liegt außerdem für jede Substanz verschieden, je nach Größe und Kompliziertheit der Baugruppen, deren Assoziationsgrad oder Mischungsverhältnis mit anderen Komponenten. Generell ist die Glasbildungstendenz hoch, wenn die Schmelze schon bei T_F eine hohe Viskosität hat, das gilt z. B. für anorganische *Netzwerkbildner*, oder wenn, wie bei vielen organischen Glasbildnern, die zunächst geringe Viskosität beim Abkühlen steil zunimmt. Glasbildung wird auch begünstigt durch vorheriges starkes Erhitzen der Schmelze, da dadurch als Kristallisationskeime wirkende Schmelzrelikte zerstört werden. Das in der Praxis übliche Vorgehen, der Schmelze des Grundglases viele weitere Stoffe zuzusetzen, erleichtert die Glasbildung teils durch die Erniedrigung der Liquidustemperatur, teils durch die zunehmende Behinderung der Ausscheidung definierter kristalliner Phasen aus der Schmelze.

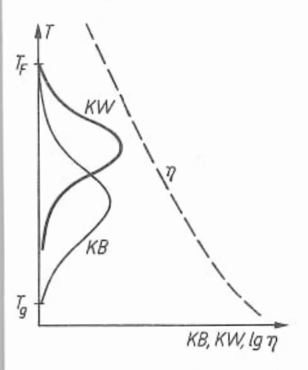


Abb. 6.4.6-3 Temperaturabhängigkeit (schematisch) der Keimbildungsgeschwindigkeit *KB*, Kristallwachstumsgeschwindigkeit *KW* und Viskosität η bei unterkühlten Schmelzen; T_F Schmelztemperatur, T_g Transformationstemperatur

Als Folge der verhinderten Kristallisation geht der jeweilige Stoff *stetig* aus dem Zustand l der Schmelze in den der unterkühlten Schmelze l_u und schließlich des Glases v über (↑ Abb. 6.4.6-2). Eine Schmelz- oder Erstarrungswärme $\Delta_F H$, wie sie allen reinen kristallinen Stoffen für die Temperatur T_F zuzuordnen ist, wird dabei nicht umgesetzt. Mit dem Auftreten des Glaszustands ist folglich keine wirklich neue Phase, sondern nur ein äußerlich anderer, anscheinend fester Zustand erreicht worden; auch die Zahl der Freiheitsgrade des Systems (↑ 6.4.5.1.) hat sich nicht verändert. Das heißt aber nicht, daß der Übergang Glas/unterkühlte Schmelze ohne jeden thermischen Effekt abliefe. Abb. 6.4.6-4 zeigt für ein ausgewähltes infrarotdurchlässiges *Chalkogenidglas* aus dem System Thallium–Germanium–Schwefel, daß durch *Differenzthermoanalyse* bei der Transformationstemperatur T_g eine endotherme Abweichung von der Nullinie registriert wird. Unterhalb T_g werden die im Glas vorliegenden Baugruppen vor allem zu thermischen Schwingungen angeregt; oberhalb ergibt sich eine Zunahme der Wärmekapazität, die mit dem „Auftauen" des Glases, also dem Entstehen fließfähiger Struktureinheiten, verknüpft ist. Da aber im Unterschied zu kristallinen Stoffen diese Änderung nicht unstetig ist, läßt sie sich nicht zwanglos in die von EHRENFEST eingeführte Ordnung der Festkörperumwandlungen einreihen.

Im allgemeinen kommt Gläsern bei der Temperatur T_g eine dynamische Viskosität η von

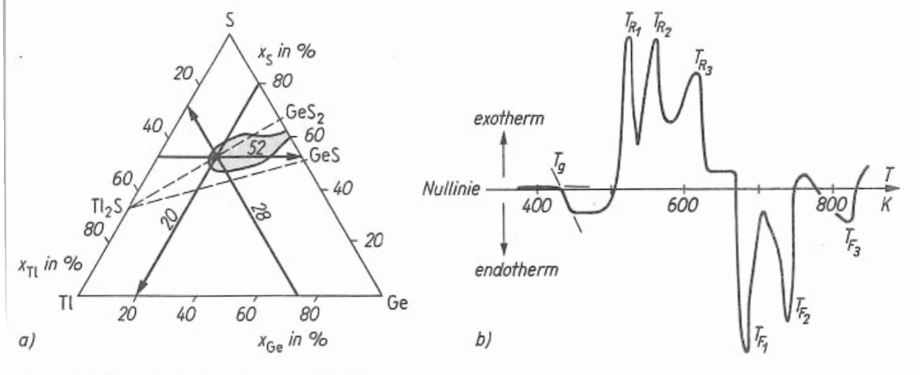


Abb. 6.4.6-4 a) Gebiet der Glasbildung (*schraffiert*) im System Thallium-Germanium-Schwefel für eine Kühlgeschwindigkeit der Schmelzen von etwa 5 Ks^{-1} im Transformationsbereich; Zusammensetzung: Molenbrüche x_i in %; b) DTA-Diagramm des in a) in Dreieckskoordinaten dargestellten Glases mit Transformationstemperatur T_g, drei Kristallisations- und drei Schmelzpeaks T_R bzw. T_F

$10^{11,5}$ bis $10^{12,5}$ Pa · s zu. Solche hohen Viskositäten werden erst weit unterhalb der Schmelztemperaturen T_F erreicht, und zwar entsprechend der Regel von KAUZMANN und BEAMAN etwa bei $T_g = aT_F$ mit $a = 0,5$ bis 0,67. Zur Überführung eines Stoffs in den glasigen Zustand ist also die Schmelze so rasch abzukühlen, daß es im Intervall $T_F \ldots T_g$ nicht zu nennenswerter Keimbildung und Kristallisation kommt. Das ist leicht bei Stoffen mit hochvernetzten Baugruppen und gerichteten Bindungen, wie bei SiO_2, GeO_2, B_2O_3, BeF_2, As_2S_3, Se, P_2O_5 und stark vernetzten organischen Polymeren. Die für eine Kristallisation erforderlichen Platzwechselvorgänge laufen hier erst nach Überwindung beträchtlicher Aktivierungsenergien ab.

Kleine Moleküle kommen leichter zur Kristallisation, besonders wenn sie hochsymmetrisch gebaut sind, da dann viele zufällige Anordnungen in der auf T_F temperierten Schmelze für ein Weiterwachsen des gebildeten Keims günstig sind. So läßt sich z. B. *Benzen* schwerer glasig erhalten als *Toluen*, das wegen der eingefügten Methylgruppe eine geringere Molekülsymmetrie aufweist.

In *Wasser* und wäßrigen Lösungen werden die an sich kleinen Moleküle durch Wasserstoffbrückenbindung (↑ 6.3.2.1.2.) assoziiert, wie die hohe Schmelztemperatur von H_2O gegenüber T_F der weniger polaren Homologen H_2S, H_2Se und H_2Te zeigt. Daraus erklärt sich die gewisse Glasbildungsneigung auch bei Wasser. Leichter als aus der Schmelze läßt es sich aus dem Dampf auf stark gekühlter Unterlage in nichtkristalliner Form abscheiden; T_g liegt hier bei etwa 100 K.

Geschmolzene Metalle und Salze kristallisieren im allgemeinen extrem leicht, da in den Schmelzen überwiegend monomere Teilchen mit weitgehender Bindungsdelokalisierung vorliegen. Nach theoretischen Abschätzungen sind hier zur Glasbildung je nach Material Kühlgeschwindigkeiten von 10^3 bis 10^{15} Ks^{-1} notwendig. Man erreicht gegenwärtig etwa 10^{10} Ks^{-1} und zwar dadurch, daß man die Schmelze in Form winziger Tröpfchen oder als Folien, Bänder und Fäden erstarren läßt, die nur wenige Nanometer, höchstens Bruchteile von Millimetern dick sind. Für die rasche Kühlung sorgen dabei z. B. mit hoher Geschwindigkeit rotierende und mit flüssigem Stickstoff gekühlte Metallwalzen. Es gelingt damit, eine große Anzahl metallischer Legierungen glasig zu erhalten, oft mit gewissen Nichtmetallanteilen, zum Beispiel bei $Pd_{72}Si_{16}Co_{12}$ oder $Pd_{43}Ni_{37}P_{20}$, in seltenen Fällen selbst reine Metalle wie Nickel, dann allerdings nur in extrem dünnen Folien.

Metallische Gläser, richtiger glasig erstarrte Metalle, weisen eine spezifische elektrische Leitfähigkeit auf, die nur wenig unter der ihrer kri-

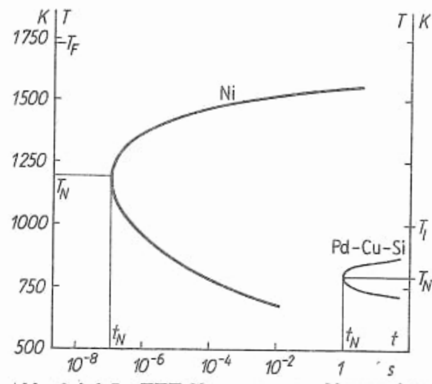


Abb. 6.4.6-5 TTT-Kurven zur Veranschaulichung des glasigen Erstarrens einer Nickel- bzw. einer Palladium-Kupfer-Silicium-Schmelze bei kritischen Kühlgeschwindigkeiten von etwa $5 \cdot 10^9$ bzw. 200 Ks^{-1} (Erläuterung im Text)

stallinen Gegenstücke liegt. Sie haben wegen des Fehlens von Korngrenzen und anderen festigkeitsmindernden Baufehlern (↑ 6.4.7.2.) eine hervorragende mechanische Stabilität, sofern sie unterhalb ihrer Transformations- oder Kristallisationstemperatur eingesetzt werden. Besonders aussichtsreich ist die Anwendung metallischer Gläser aus ferromagnetischen Legierungen, da hier die Vorzüge ausgeprägt weichmagnetischer Materialien mit einer ungewöhnlich hohen mechanischen Stabilität zusammentreffen.

Die bei der Herstellung von Gläsern zu beachtenden kinetischen Bedingungen lassen sich in einem TTT-Diagramm (↑ 6.4.5.5.) veranschaulichen, das die Kühlgeschwindigkeit angibt, bei der der Übergang Schmelze – Kristall zugunsten des glasigen Erstarrens ausbleibt (Abb. 6.4.6-5). Die Kurven sind für einen zulässigen Kristallitanteil am Glasvolumen von $10^{-4}\%$ ermittelt. Sie geben zu jeder Temperatur T die Zeit t an, bei der gerade dieser Grad an Kristallinität erhalten wird; damit ist zugleich die *kritische Mindestkühlgeschwindigkeit* $(dT/dt)_{kr}$ für das glasige Erstarren festgelegt. Näherungsweise gilt $(dT/dt)_{kr} \approx (T_F - T_N)/t_N$; T_N und t_N sind dabei die Koordinaten der „Kurvennase", T_F die Schmelz- oder Liquidustemperatur. Wegen des Näherungscharakters der hierfür zugrunde gelegten Gleichungen der Transformationskinetik sind deren Aussagen aber möglichst durch unabhängige Messungen zu stützen.

6.4.7. Realstruktur von Kristallen

6.4.7.1. Einteilungsprinzip. Der ideale Gitteraufbau, gekennzeichnet durch die strenge drei-

dimensionale Periodizität in der Anordnung der Bausteine (↑ 6.4.1.), läßt sich unter normalen Druck- und Temperaturbedingungen nicht realisieren. Die Untersuchung der statt dessen vorliegenden Realstruktur von Kristallen stellt heute eine wesentliche Forschungsrichtung der Kristallphysik und -chemie dar, da die weitaus meisten Kristalleigenschaften, chemische, physikalisch-chemische, elektrische, magnetische, mechanische u. a. durch die Realstruktur bestimmt oder zumindest beeinflußt werden. Es ist zweckmäßig, von struktur- bzw. störungsempfindlichen und von störungsunempfindlichen Eigenschaften zu sprechen.

Unter der *chemischen Fehlordnung* versteht man die Abweichungen vom Idealkristall durch den Einfluß von Beimengungen bzw. Dotierungen. Die *strukturelle Fehlordnung* resultiert aus den geometrischen Abweichungen der Gitterbausteine von der regelmäßigen Anordnung im Idealgitter. Daneben kann man eine *elektrische Fehlordnung* als Abweichung der Verteilung der elektrischen Ladung von der Periodizität definieren; sie ist im allgemeinen an die chemische und strukturelle Fehlordnung gebunden und somit sekundärer Natur. Die Behandlung der Fehlordnungsarten ist z. B. Grundlage für das Verständnis von Diffusion, Kristallwachstum, Festkörperreaktivität und -aktivierung.

Hinsichtlich der Dimensionen der Wirkungsbereiche von Baufehlern werden nach KLEBER folgende Gruppen unterschieden:

Bei *makroskopischen Baufehlern* betragen die linearen Abmessungen mehr als 10^{-3} cm. Hierher gehören grobe Störungen, wie Sprünge, Risse, Einlagerungen und dergleichen.

Die lineare Ausdehnung *mikroskopischer Baufehler* liegt zwischen 10^{-5} und 10^{-3} cm.

Submikroskopische Baufehler haben Abmessungen unterhalb 10^{-5} cm.

Eine weitere Untergliederung ist möglich in ultramikroskopische, amikroskopische und atomare Baufehler.

6.4.7.2. Chemische Fehlordnung

6.4.7.2.1. *Mischkristallbildung.* Eine besondere Art chemischer Fehlordnung tritt bei den *Mischkristallen* auf (↑ 6.4.4.3.). Der einfachste Fall der Bildung von *Substitutionsmischkristallen* liegt vor, wenn eine Gastkomponente, z. B. Kaliumbromid, so in einen Wirtskristall, z. B. Kaliumchlorid, eingebaut wird, daß die Atome der Fremdsubstanz auf normale Gitterplätze zu liegen kommen. Beide Komponenten haben in diesem Fall ähnliche Atom- bzw. Ionenradien (↑ Tab. 6.4.4-2) und meist einen ähnlichen Ladungszustand ihrer Teilchen. Eine Substitution ist aber auch dann möglich, wenn die Komponenten bei ähnlicher Größe unterschiedliche Wertigkeit besitzen, z. B. Ca^{2+}/Na^{+} oder Ca^{2+}/Al^{3+}. In diesem Fall kann der Austausch der Kationen nicht direkt erfolgen; es müssen wegen der Elektroneutralitätsbedingung weitere Veränderungen zur Kompensation des Ladungsunterschieds stattfinden. Man spricht hier von einer *gekoppelten Substitution*.

Sind im Kristall Kationen oder Anionen im Über- bzw. Unterschuß gegenüber der stöchiometrischen Formel vorhanden, können die fehlenden Ladungsträger durch Elektronen ersetzt werden.

Werden bestimmte Kristalleigenschaften durch chemische Fehlordnung vorteilhaft verändert, so strebt man Abweichungen von der Stöchiometrie bei der Kristallzüchtung durch bestimmte Zusätze, die *Dotierungen*, bewußt an (↑ 6.4.2.1.).

6.4.7.2.2. *Farbzentren in Ionenkristallen.* Sie bilden sich in den sonst farblosen Alkalihalogenidkristallen durch die sogenannte *additive Verfärbung* beim Tempern der Kristalle im entsprechenden Alkalimetalldampf nahe der Schmelztemperatur. Dabei werden Elektronen in den Kristall injiziert. Die Kompensation des negativen Ladungsüberschusses erfolgt durch Ausscheidung von Anionen aus dem Gitterverband auf dem Wege der Leerstellendiffusion (↑ 6.4.8.2.).

Die *subtraktive Verfärbung* erfolgt unter Einwirkung energiereicher Röntgen- oder γ-Strahlen bei normalen oder tiefen Temperaturen. Neben Farbzentren werden dabei auch andere Gitterdefekte gebildet.

Die *fotochemische Verfärbung* setzt eine Sensibilisierung des Kristalls bei seiner Züchtung voraus; bei Kaliumchlorid gelingt das z. B. durch Dotieren mit Kaliumhydrid.

Die Farbzentren weisen unterschiedliche Struktur auf. Das einfachste Farbzentrum ist das *F-Zentrum* (Abb. 6.4.7-1), wie es vorrangig beim Abschrecken additiv verfärbter Kristalle

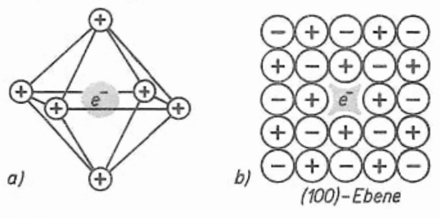


Abb. 6.4.7-1 Strukturmodell des F-Zentrums nach PICK; a) räumliche Anordnung, b) in der (100)-Ebene

auf tiefe Temperaturen entsteht. Ein Halogenidion wird durch Diffusionsprozesse aus dem Gitterverband verdrängt und zum Ausgleich der fehlenden negativen Ladung durch ein Elektron ersetzt. Dadurch verändern sich die Absorptionseigenschaften charakteristisch.

Durch spezielle fotochemische Reaktionen können aus *F*-Zentren kompliziertere, z. B. die *F'*-, *M*-, *V*-, *N*- und *R*-Zentren gebildet werden.

Das *F'-Zentrum* entsteht unter Lichteinstrahlung aus zwei *F*-Zentren unter Bildung je einer mit zwei Elektronen besetzten oder unbesetzten Anionenleerstelle. Ein *R-Zentrum* ist ein Agglomerat dreier *F*-Zentren.

6.4.7.2.3. *Überschußleitung in Verbindungshalbleitern.* Die Nichtstöchiometrie bestimmter Kristalle kann von einem Über- oder Unterschuß an Kationen bzw. Anionen herrühren, wobei für den Ladungsausgleich verschiedene Möglichkeiten bestehen. In Zinkoxid- und Cadmiumoxidkristallen, ZnO bzw. CdO, liegt ein *Kationenüberschuß* vor, der durch Sauerstoffabgabe bei erhöhter Temperatur zustandekommt. Der äquivalente Anteil von Elektronen und Zink(II)-ionen bleibt auf Zwischengitterplätzen; in anderen Fällen kommt es auch vor, daß die Ionen auf Gitterplätzen verbleiben und Elektronen in Fehlstellen eingebaut werden. Da die negativ geladenen Elektronen, die den Kationenüberschuß kompensieren, bei Anlegen eines äußeren Felds den Hauptanteil zur Leitfähigkeit beitragen, bezeichnet man diesen Typ als *n-Leiter*. Weitere Beispiele für Metallüberschußhalbleiter sind BeO, BaO, U_3O_8, WO_3, CdS, SnO_2 und MoO_3.

Kationenunterschuß liegt z. B. bei Nickeloxid und Kupfer(I)-oxid, NiO bzw. Cu_2O, vor. Hier wird die Ladung der fehlenden Kationen dadurch kompensiert, daß die entsprechende Zahl Metallionen in einen höheren Ladungszustand übergeht: $Ni^{2+} \rightarrow Ni^{3+} + e^-$ und $Cu^+ \rightarrow Cu^{2+} + e^-$. Die Leerstelle □ eines Ions Ni^{2+} kann somit durch zwei Ni^{3+} kompensiert werden (Abb. 6.4.7-2). Die Orte der höhergeladenen Kationen, die ein zusätzliches Elektron abgegeben haben, bezeichnet man als Elektronendefektstellen oder *Löcher*. Sie bewirken ebenfalls bei Anlegen eines äußeren Felds eine elektrische Leitfähigkeit; man spricht hier von *p-Leitern*. Weitere Metallunterschußhalbleiter sind FeO, FeS, CoO, Cu_2S, Ag_2O und $ZnCr_2O_4$.

O^{2-} Ni^{2+} O^{2-} Ni^{3+} O^{2-}
Ni^{2+} O^{2-} Ni^{2+} O^{2-} Ni^{2+}
O^{2-} Ni^{3+} O^{2-} □ O^{2-}
Ni^{2+} O^{2-} Ni^{2+} O^{2-} Ni^{2+}

Abb. 6.4.7-2 NiO als Metallunterschußhalbleiter

Ein besonderes Verhalten zeigt Titanium(II)-oxid, TiO, welches Steinsalzstruktur aufweist (↑ Tab. 6.4.4-7) und in seiner Zusammensetzung von $TiO_{1,35}$ bis $TiO_{0,60}$ schwanken kann. Je nach dem O/Ti-Verhältnis ist es *n*- oder *p*-leitend.

Auch ohne Kationenleerstellen und ohne Veränderungen der Anionenzahl lassen sich Halbleitereigenschaften erzeugen, z. B. in lithiumdotierten NiO-Kristallen. Die Ladung der Ionen Li^+ wird durch Ni^{3+}-Ionen, alle auf normalen Kationplätzen, kompensiert, so daß ein Paar Ni^{3+}/Li^+ gerade ein Paar Ni^{2+}/Ni^{2+} substituiert. Die Zusammensetzung kann somit durch die Formel $\overset{+1}{Li}_x\overset{+3}{Ni}_x\overset{+2}{Ni}_{1-2x}\overset{-2}{O}_1$ ausgedrückt werden. Im Unterschied zu reinem Nickel(II)-oxid, das blaßgrün und nichtleitend ist, ist das mit Lithium dotierte Material schwarz und *p*-leitend; bei $x = 0{,}1$ hat es z. B. eine um den Faktor 10^8 größere Leitfähigkeit als reines NiO.

6.4.7.2.4. *Dotierte Elementhalbleiter. Germanium* und *Silicium*, Elemente der IV. Hauptgruppe des Periodensystems, besitzen in reinem Zustand eine nur geringe Leitfähigkeit. Sie läßt

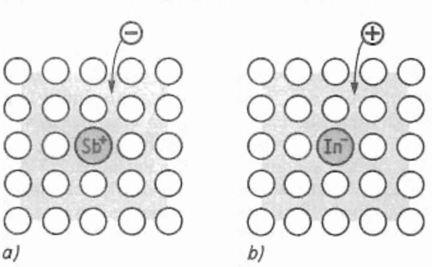


Abb. 6.4.7-3 a) durch Substitution von Germanium mit Antimon entstandene Donatorstörstelle mit negativer Ladungswolke, b) durch Substitution von Germanium mit Indium entstandene Akzeptorstörstelle mit positiver Ladungswolke

sich durch Einbringen elektrisch aktiver Störstellen stark erhöhen. Durch Einbau von Elementen der V. Hauptgruppe (Phosphor, Arsen oder Antimon) werden Störstellen vom Donatortyp erzeugt, da das überschüssige 5. Valenzelektron, eins je substituiertes Atom, nur leicht gebunden ist und in der Umgebung des Defekts eine ausgedehnte, bei Germanium etwa 1000 Atome umfassende, negative Ladungswolke erzeugt (Abb. 6.4.7-3a). Da Elektronen die Ladungsträger sind, handelt es sich hier um *n-Leiter*.

Die Substitution von Germanium und Silicium durch Bor, Aluminium, Indium oder Gallium, also Elemente der III. Gruppe, erzeugt Störstellen vom Akzeptortyp (Abb. 6.4.7-3b). Das ursprünglich neutrale Dotierungselement, z. B. In, nimmt leicht ein Elektron auf, so daß eine positive Ladungswolke um das Teilchen In^- entsteht. Die somit im Wirtskristall erzeugten positiven Löcher können wie die Elektronen bei der *n*-Leitung thermisch angeregt werden und zur elektrischen Leitung beitragen. Man spricht hier von *p-Leitern* (s. a. 6.4.7.2.3.).

6.4.7.3. Strukturelle Fehlordnung. Die Baufehler in Kristallgittern lassen sich nach ihrer

räumlichen Ausdehnung einteilen in Punkt-, Linien- und Flächen- bzw. null-, ein- und zweidimensionale Defekte. Neben den nachstehend behandelten Defekten gibt es noch weitere komplizierte, wenig definierte Baufehler, z. B. Anhäufungen von Punktdefekten zu „Clustern" [*cluster* engl., Schwarm], die atomare Dimensionen überschreiten. Die verschiedenen Baufehler treten meist nicht in reiner Form im Kristall auf; sie beeinflussen sich gegenseitig und können unter- und miteinander reagieren.

6.4.7.3.1. *Punktdefekte.* Die vier Grundtypen atomarer Fehlordnung sind in Abb. 6.4.7-4 für binäre Ionenkristalle vom Typ *AB* zusammengestellt. Bei den Typen nach Frenkel sind entweder die Kationen *A* oder die Anionen *B* fehlgeordnet, bei den Typen nach Schottky beide Teilchensorten zugleich, entweder durch Leerstellenbildung in beiden Teilgittern *A*, *B* oder durch Einnehmen von Zwischengitterplätzen im ansonsten voll besetzten Gitter. Im allgemeinen sind die Fehlordnungstypen nicht rein verwirklicht; bei den Alkalimetallhalogeniden z. B. überwiegt aber die Schottkyfehlordnung.

Punktdefekte unterscheiden sich von den Linien- und Flächendefekten grundsätzlich dadurch, daß sie im thermischen Gleichgewicht sind. Es ist also prinzipiell nicht möglich, Idealkristalle bei einer von 0 K abweichenden Temperatur zu realisieren, da stets mit der entsprechenden Gleichgewichtskonzentration atomarer Defekte zu rechnen ist. Daher bezeichnet man das Auftreten von Punktdefekten auch als *Eigenfehlordnung* oder thermische bzw. *thermodynamische Fehlordnung.*

Fehlstellen im Gitter bewirken einen Anstieg sowohl der inneren Energie *U* als auch der

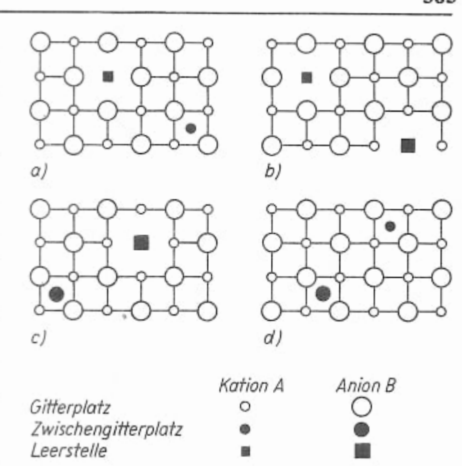


Abb. 6.4.7-4 Die vier Grundtypen der Punktfehlordnung in binären Ionenkristallen *AB*; a) Frenkelfehlordnung, b) Anti-Frenkel-Fehlordnung, c) Schottkyfehlordnung, d) Anti-Schottky-Fehlordnung

6.4.7.3.2. *Liniendefekte.* Man unterscheidet hier zwischen den Grenzfällen einer *Stufen-* bzw. *Schraubenversetzung* (Tafel 38). Unter der *Versetzung* oder Versetzungslinie allgemein versteht man die Grenzlinie zwischen ungestörtem und gestörtem Kristallbereich. Exakt kann eine Versetzung durch den Umlauf nach Burgers beschrieben werden, indem man die Versetzungslinie in einer zu ihr senkrecht liegenden Ebene umfährt. Die gegenüber dem Idealgitter

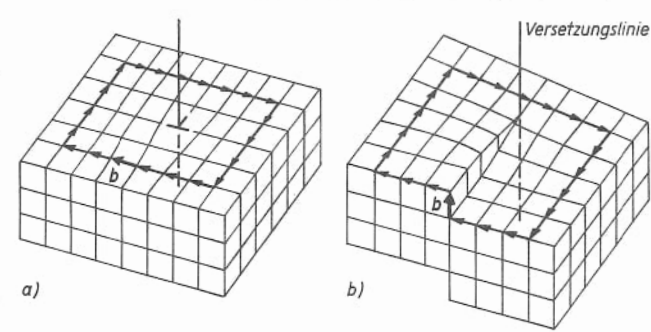


Abb. 6.4.7-5 Burgersumlauf a) bei einer Stufenversetzung, b) bei einer Schraubenversetzung

Entropie *S* des Systems. Die Gleichgewichtskonzentration an Fehlstellen läßt sich folglich über die Bestimmung von ΔU und ΔS und deren Verknüpfung zu $\Delta F = \Delta U - T\Delta S = 0$ ermitteln (s. a. 6.1.2.4.). Als molare Energie der Bildung von Schottkydefektpaaren werden bei Alkalimetallhalogeniden Werte von 100 bis 300 kJmol^{-1} erhalten; das ist etwa ein Viertel der Gitterenergie (↑ 6.4.3.1.).

beobachtete Verschiebung der Atome wird durch den Burgersvektor ***b*** charakterisiert; ***b*** ist bei den sogenannten vollständigen Versetzungen ein Gittervektor im Bravaisgitter der betrachteten Struktur.

Bei der Stufenversetzung (Abb. 6.4.7-5a) gelangt man während des Umlaufs zum Ausgangspunkt zurück, ohne die Ebene zu verlassen; der Burgersvektor verläuft in diesem Fall senk-

recht zur Versetzungslinie. Die Stufenversetzung wird mit dem Symbol ⊥ belegt, wobei der senkrechte Strich die eingeschobene Netzebene und der durchgehende waagerechte Strich die Gleitrichtung angeben.
Bei der Schraubenversetzung (Abb. 6.4.7-5b) verläuft der Burgersvektor, um den der Umlauf den Ausgangspunkt verfehlt, parallel zur Versetzungslinie.

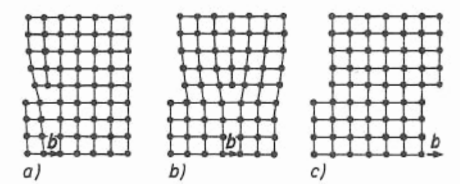


Abb. 6.4.7-6 Entstehen und Wandern einer Stufenversetzung

Im allgemeinen ist die Versetzungslinie gekrümmt, und man findet gleichzeitig Stufen- und Schraubenanteile. In jedem Fall führt aber eine Versetzung zur Anhäufung zusätzlicher Energie im Kristall, da sich die beteiligten Atome in verzerrten Lagen befinden.
Versetzungen können durch mechanische Beanspruchung eines Kristalls entstehen. In Abb. 6.4.7-6a ist das Abgleiten des oberen Teilgitters dargestellt; unter der Wirkung einer Schubspannung bewegt sich die Versetzung durch den Kristall und verläßt ihn schließlich wieder (Abb. 6.4.7-6b, c). Dadurch entsteht an der Oberfläche eine Elementarstufe, deren Höhe der Betrag des Burgersvektors ist. Je nach der Zahl der gewanderten Versetzungen in einer Gleitebene kann die entstehende Gleitstufe ein Vielfaches davon betragen (s. a. 6.4.8.1.3.).
Versetzungen können im Innern eines perfekten Kristalls weder beginnen noch enden, sondern müssen in sich selbst zurückführen unter Bildung eines *Versetzungsrings*; oder sie durchstoßen die Oberfläche unter Bildung einer *Versetzungsschleife*. In einer bestimmten Struktur sind nur wenige Burgersvektoren und damit Versetzungsformen stabil. Der Grund liegt in den Versetzungsreaktionen, in deren Verlauf sich instabile Burgersvektoren unter Energieabgabe in stabile umbilden. Die letztlich verbleibenden Gleitvektoren entsprechen im allgemeinen bei vollständigen Versetzungen kürzesten Gitterabständen im Bravaisgitter. Je Zentimeter Versetzung ergeben sich bei Metall- und Ionenkristallen Energien der Größenordnung 10^{-11} bis 10^{-10} J.
Versetzungen können in einem versetzungsfreien Kristall durch homogene Erwärmung, d. h. durch rein thermische Bewegungen der Bausteine, nicht entstehen. Die dazu notwendige Aktivierungsenergie kann erst durch Mitwirkung einer äußeren Spannung aufgebracht werden. Enthält ein Kristall bereits Versetzungen, z. B. wachstumsbedingte, so genügen bereits äußerst kleine Spannungen, um die Versetzungsdichte durch sogenannte Multiplikationsprozesse zu erhöhen. Die *Versetzungsdichte* ist die Zahl der Versetzungen, die je Einheitsfläche 1 cm^2 austreten; sie kann zwischen 0 und 10^{12} cm^{-2} schwanken und ist von Herstellung und Nachbehandlung der Kristalle abhängig. Man ist heute schon in der Lage, von bestimmten Stoffen, z. B. von Silicium und Germanium, versetzungsfreie Kristalle größerer Abmessungen herzustellen.

6.4.7.3.3. *Überlagerung von Punkt- und Liniendefekten.* Die behandelten Punkt- und Liniendefekte sind nicht unabhängig voneinander. Weist z. B. eine durch Netzebeneneinschub entstandene Stufenversetzung einen Sprung auf (Abb. 6.4.7-7), können im Gitter durch Anlagern von Atomen Leerstellen erzeugt oder durch Abbau der Stufe Zwischengitteratome geschaffen werden. Dadurch kommt es zu einem Klettern der Versetzung senkrecht zur Gleitrichtung. Solche Wechselwirkungen sind besonders bei hohen Temperaturen zu berücksichtigen, wo die Platzwechselvorgänge erleichtert sind. Versetzungen können auch durch Kondensation von Leerstellen im Gitter entstehen. Abb. 6.4.7-8 zeigt einen auf diese Weise gebildeten *Versetzungsring*. Versetzungsringe können in Kristallen entstehen, wenn diese von hohen

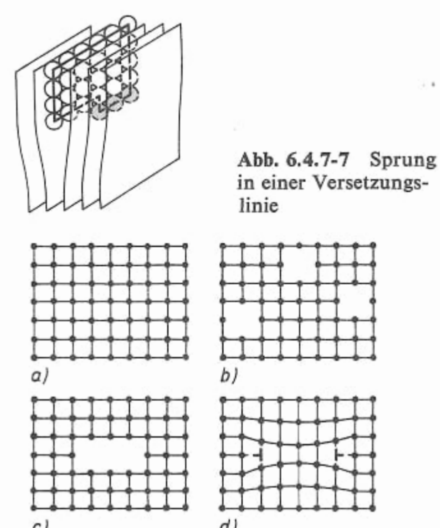


Abb. 6.4.7-7 Sprung in einer Versetzungslinie

Abb. 6.4.7-8 a) Idealgitter, b) isolierte Leerstellen, c) Leerstellenkondensation, d) Entstehen des Versetzungsrings

Temperaturen abgeschreckt werden, wobei Leerstellen einfrieren und kondensieren. Umgekehrt können Leerstellen durch bestimmte Versetzungsreaktionen erzeugt werden, so in Abb. 6.4.7-9 durch gegenseitiges Aufheben zweier Versetzungen; räumlich gesehen, wird dabei eine ganze Leerstellenreihe längs der Versetzungslinie erzeugt.

6.4.7.3.4. *Flächendefekte.* Die wichtigsten zweidimensionalen Baufehler sind Korngrenzen, Stapelfehler und Zwillinge. Versetzungen können spezielle Anordnungen in einem Kristall bilden und Korngrenzen aufbauen. Durch eine *Korngrenze* werden Kristallgebiete unterschiedlicher Orientierung getrennt. Von einer *Kleinwinkelkorngrenze* spricht man, wenn die aneinanderstoßenden Kristallgebiete nur einige Winkelsekunden bis -minuten, im Extremfall 10 bis 20 Grad, gegeneinander geneigt sind. Im einfachsten Fall ist eine Korngrenze aus einer Versetzungsart aufgebaut, z. B. nur aus Stufenversetzungen (Abb. 6.4.7-10). Die beiden aneinandergrenzenden Kristallgebiete sind auf beiden Seiten der Korngrenze gegeneinander geneigt, wobei die Achse in der Korngrenze liegt. Ist ϑ der Neigungswinkel, D der Abstand zwischen den Durchstoßpunkten von Versetzungen und b der Burgersvektor, so läßt sich für kleine Neigungswinkel die Beziehung $\vartheta = b/D$ ableiten. Sie gilt lediglich für die Sonderfälle, daß die Kleinwinkelkorngrenzen nur aus Stufen- oder nur aus Schraubenversetzungen aufgebaut sind. Der Versetzungsabstand D kann einer ge-

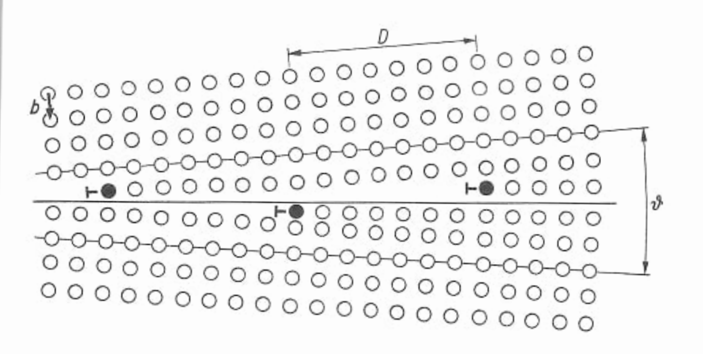


Abb. 6.4.7-10 Aus Stufenversetzungen aufgebaute Kleinwinkelkorngrenze (Erläuterung im Text)

ätzten Oberfläche mit Kleinwinkelkorngrenze entnommen werden, da sich unter günstigen Voraussetzungen den Ätzgruben Versetzungsdurchstoßpunkte zuordnen lassen. Beim Aufbau einer Kleinwinkelkorngrenze aus reinen Schraubenversetzungen sind die angrenzenden Subkörner gegeneinander verschränkt, wobei die Drehachse senkrecht zur Korngrenze steht. Auf *Großwinkelkorngrenzen* sind einfache Versetzungsmodelle nicht mehr anwendbar, da der Orientierungswinkel zu groß ist. Man nimmt an, daß in der sogenannten *Beilbyschicht* die Kristalle derart fehlgeordnet sind, daß man von einer flüssigkeitsanalogen Struktur sprechen kann; es gibt dann keine kristallographisch definierte Übergangsschicht mehr (Abb. 6.4.7-11 a). Durch neuere Experimente wurde das Modell gemäß Abb. 6.4.7-11 b als wahrscheinlich angenommen. Danach ist die Dicke der Korngrenze – mit ein bis zwei Gitterabständen –

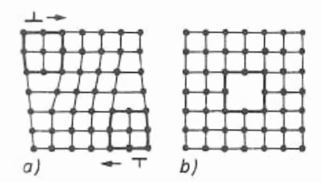
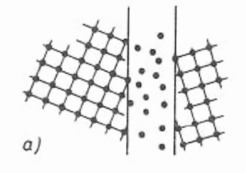


Abb. 6.4.7-9 Erzeugen einer Leerstellenreihe durch Annihilation zweier ungleichnamiger Versetzungen

wesentlich geringer; die Korngrenze enthält einige Atome A, die beiden Gittern gleichzeitig, einige Atome B, die keinem von beiden angehören. Dieses Korngrenzenmodell hat einen höheren Ordnungsgrad als das Modell nach BEILBY. Kristalle mit Korngrenzen sind gegenüber Einkristallen um die spezifische Korngrenzenenergie energiereicher; deren Werte liegen für Metalle bei 10^{-5} bis 10^{-4} J je cm² Grenzfläche.

Die überwiegend bei Metallkristallen und bei



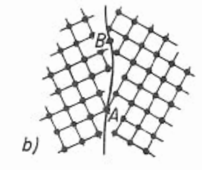


Abb. 6.4.7-11 Korngrenzenmodelle nach FRIEDEL; a) Beilbyschicht, b) wahrscheinlicher Aufbau

einigen Schichtengittern beobachteten *Stapelfehler* sind Störungen in der normalen Schichtenfolge beim Aufbau einer Struktur, z. B. von dichtesten Kugelpackungen (↑ 6.4.4.1.). Hierbei bleibt die Zahl der nächsten Nachbarn eines herausgegriffenen Atoms, d. h. dessen Koordinationszahl, konstant. Nur die gegenseitige Anordnung der Nachbarn in zweiter und höherer Koordination weicht vom Idealzustand ab. Daher sind die energetischen Unterschiede Stapelfehler enthaltender Strukturen äußerst gering und experimentell schwer zugänglich.

Zwillinge können auf verschiedene Weise entstehen, z. B. als Wachstumszwillinge oder durch mechanische Beanspruchung als Deformationszwillinge. Im einfachsten Fall kann man sich den Aufbau eines Zwillings durch Spiegelung eines Kristallbereichs an einer *Zwillingsebene*, die im allgemeinen eine einfach indizierte Fläche ist, verständlich machen (Abb. 6.4.7-12).

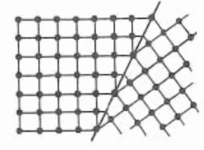


Abb. 6.4.7-12 Kohärenter Zwilling mit Spiegelung an der Zwillingsebene (112)

6.4.8. Kristallsymmetrie und gitterphysikalische Eigenschaften

Aus der Definition des kristallinen Zustands (↑ 6.4.1.) folgt, daß viele Eigenschaften von Kristallen, z. B. die mechanischen, optischen und elektrischen, im allgemeinen richtungsabhängig, also vektorieller Natur sind. Für ausgewählte mechanische Eigenschaften sowie für Diffusion und thermische Ausdehnung soll nun deren Verknüpfung mit der Kristallsymmetrie aufgezeigt werden.

6.4.8.1. Mechanische Eigenschaften

6.4.8.1.1. *Elastische Deformation.* Beansprucht man einen Kristall durch Zug oder Druck mechanisch, wird er deformiert. Ist die Verformung reversibel, so spricht man von elastischer Deformation.

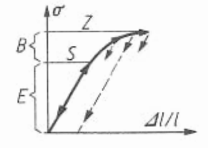


Abb. 6.4.8-1 Spannungs-Dehnungs-Diagramm (schematisch); *B* Bereich der bleibenden plastischen Verformung, *E* Elastizitätsbereich

Das Deformationsverhalten wird durch das *Spannungs-Dehnungs-Diagramm* beschrieben, in dem die Spannung σ gegen die Dehnung, das ist die relative Längenänderung $\Delta l/l$ der Probe, aufgetragen wird. Bei geringer Belastung sind Normalspannung σ und Längenänderung proportional zueinander (Abb. 6.4.8-1). Der Festkörper wird gemäß dem Gesetz von Hooke elastisch deformiert: $\sigma = E\Delta l/l$. Der Proportionalitätsfaktor E ist der *Elastizitätsmodul* oder *Youngsche Modul*; er hat die Größenart eines Drucks. Für Schubspannungen τ nimmt das Hookesche Gesetz die Form $\tau = \gamma G$ an; γ ist der Winkel, um den sich die der Schubspannung ausgesetzte Fläche gegenüber dem Ausgangszustand verschiebt, G ist der *Gleit-* oder *Schubmodul*, für dessen Verknüpfung mit E sowie den weiteren elastischen Konstanten, dem *Kompressionsmodul K* und der *Poissonschen Konstanten* μ, einfache Beziehungen gelten.

Wegen der Anisotropie im Kristallaufbau ist die elastische Deformierbarkeit der Kristalle in den verschiedenen kristallographischen Richtungen verschieden. Für den Elastizitätsmodul von Eisen ergeben sich z. B. für Beanspruchungen in den Kristallrichtungen [100], [110] bzw. [111] die Werte $E = 132$, 206 bzw. 284 GPa; für polykristalline Proben erhält man etwa 200 GPa.

6.4.8.1.2. *Plastische Deformation.* Sie liegt vor, wenn die Formänderung auch nach Aufheben der äußeren Einwirkung erhalten bleibt. Die *Elastizitäts-*, *Streck-* oder *Fließgrenze S* (↑ Abb. 6.4.8-1) gibt die maximale Spannung σ an, bis zu der ein Kristall elastisches Verhalten zeigt. Bei weiterer Erhöhung von σ treten bleibende Gitterdeformationen auf (Tafeln 40/41), die eine Abweichung von der Hookeschen Geraden bewirken und schließlich bei der *Zerreiß-* oder *Bruchgrenze Z* zum Zerreißen des Kristalls führen. Z bestimmt die materialspezifische Reiß- oder Bruchfestigkeit.

Im wesentlichen können die plastischen Eigenschaften der Kristalle auf zwei Elementarprozesse zurückgeführt werden, die mechanische Translation und die Zwillingsbildung durch mechanische Beanspruchung; diese lassen sich ihrerseits unter dem Begriff der Gleitung zusammenfassen.

6.4.8.1.3. *Gleitung.* Die *mechanische Translation* besteht in einer Parallelverschiebung von Schichten im Kristall längs wichtiger, d. h. einfach indizierter Flächen in kristallographisch ausgezeichneten Richtungen. Ihre beiden geometrischen Elemente bezeichnet man entsprechend als *Gleitebene* und *Gleit-* bzw. *Translationsrichtung* (Abb. 6.4.8-2). Ihre Mannigfaltigkeit wird durch den Gitterbau bestimmt. Die gute Verformbarkeit von Aluminium oder Kupfer geht z. B. darauf zurück, daß in deren kubisch-flächenzentrierten Gittern besonders

viele günstige Kombinationen von Gleitebenen und -richtungen vorkommen. Schematisch sind zwei Translationsgleitungen für den NaCl-Typ in Abb. 6.4.8-3 dargestellt. Das Abgleiten der Schichtpakete erfolgt nicht so, daß sich alle Bausteine einer Ebene gleichzeitig verschieben; es ist nur über eine Versetzungsbewegung möglich, wie sie in Abb. 6.4.7-6 beschrieben wurde. Die sukzessive Abgleitung der Gitterebenen durch Bewegung von Versetzungen macht verständlich, daß die erforderlichen Schubspannungen τ um mehrere Größenordnungen niedriger sind als die theoretische Festigkeit des Idealkristalls und stark von der Kristallvorbehandlung abhängen. Im allgemeinen erhöht sich mit zunehmender Deformation der Widerstand gegen die Verformung, d. h., es tritt mit zunehmender Versetzungsdichte eine Verfestigung ein. Trägt man den Zusammenhang zwischen Schubspannung τ und der Verformung durch Abgleiten auf, erhält man den in Abb. 6.4.8-4 dargestellten typischen Verlauf einer Verfestigungskurve. Grundsätzlich können drei Bereiche unterschieden werden, deren Ausdehnung weitgehend von der Vorbehandlung und Orientierung des Kristalls sowie von Verformungsgeschwindigkeit und -temperatur abhängt. Im Bereich I nimmt die Spannung mit zunehmender Abgleitung geringfügig zu, die plastische Verformung bewirkt kaum eine Verfestigung. Hier steigt durch Betätigung des günstig orientierten Gleitsystems die Zahl von Versetzungen aber schon stark an. In II kommt es infolge der hohen Versetzungsdichte zu Wechselwirkungen zwischen den Versetzungen und zur Betätigung weiterer Gleitsysteme. Es entstehen Versetzungen, die weder gleiten noch klettern können und somit ein Hindernis für die Bewegung benachbarter Versetzungen darstellen. Die weitere plastische Verformung setzt wesentlich höhere Spannungen τ voraus; das Material verfestigt sich stark. Im Bereich III geht der Verfestigungsanstieg zurück. Es erfolgt ein Teilabbau der Barrieren, zum Beispiel dadurch, daß Versetzungen durch Kletterbewegungen bestehende Spannungsfelder umgehen und sich an der Kristalloberfläche auflösen. Bei der Zwillingsbildung durch mechanische Einwirkung verschieben sich wie bei der Translation einzelne Kristallbereiche gegeneinander, die beiden Teile geraten aber in spiegelsymmetrische Stellung zueinander. Dieser Verformungsmechanismus tritt in seiner Bedeutung weit zurück gegenüber dem durch reine Translation.

6.4.8.1.4. *Bruch.* Der Bruch in einem Kristall wird durch zwei Teilprozessse bestimmt, durch die Entstehung eines Risses und durch seine Ausbreitung. Als bruchauslösende Mechanismen kommen verschiedene Versetzungsreaktionen in Frage. Abb. 6.4.8-5 zeigt zwei Gruppen von je vier Versetzungen entgegengesetzten Vorzeichens, die gegeneinander aufgelaufen sind. Im linken Teilbild sind durch die gekrümmten Linien die Spannungen angedeutet, die entstehen, wenn die vier Versetzungen mit dem Burgersvektor $b = 4a$ in zwei um den Abstand $h = 5a$ voneinander entfernten Gleitebenen auf-

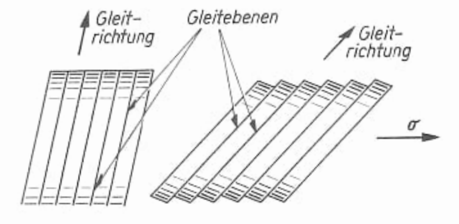


Abb. 6.4.8-2 Modell des Gleitvorgangs am Beispiel eines Einkristallstabs; σ maximale Spannung

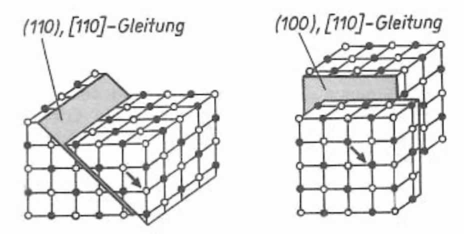


Abb. 6.4.8-3 Translationsgleitung in gleicher Richtung, jedoch verschiedenen Ebenen bei Kristallen mit NaCl-Struktur

Abb. 6.4.8-4 Typische Verfestigungskurve kubisch-flächenzentrierter Metalle

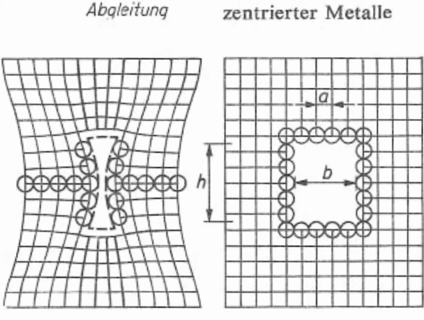


Abb. 6.4.8-5 Auflaufen zweier Gruppen von je vier Versetzungen und Entstehung eines Risses nach KOCHENDÖRFER; a Gitterkonstante, b und h Abmessungen des Risses

gelaufen sind. Es werden Normalspannungen erreicht, die die Kohäsionskräfte des Gitters überwinden. Das rechte Teilbild zeigt den entstandenen Riß der Breite *b* und Höhe *h*. Durch ihn gehen die ursprünglichen elastischen Verspannungen des Gitters zurück und treten als Oberflächenenergie des Anrisses in Erscheinung.

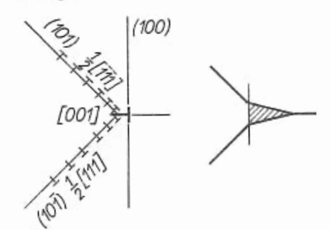
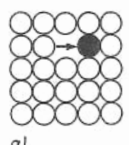


Abb. 6.4.8-6 Zwei Gruppen aufgelaufener Versetzungen bilden einen Riß in der (100)-Ebene (nach COTTRELL)

Laufen einzelne Versetzungen verschiedener Gleitebenen aufeinander, so können ebenfalls Normalspannungen entstehen, die die Gitterfestigkeit überschreiten. Abb. 6.4.8-6 zeigt zwei Gruppen von aufgelaufenen Versetzungen in den Ebenen (101) und (10$\bar{1}$), durch die ein Riß in der (100)-Ebene gebildet wird. Durch eine äußere Schubspannung ist die Versetzungsdichte am Kopf der Gruppe so groß geworden, daß die Normalspannungswerte dafür ausreichen. Die Versetzungen reagieren mit den Burgersvektoren 1/2 [111] und 1/2 [$\bar{1}$11] und bilden Versetzungen mit dem Gleitvektor [001], der in der (100)-Ebene liegt. Dadurch entsteht zuerst ein Keil minimaler Größe, bis durch weitere auflaufende Versetzungen das Gitter aufreißt. Die bei der Bildung des Anrisses freiwerdende elastische Energie genügt allein nicht, um den Riß auszubreiten, hierzu müssen unter der Wirkung der äußeren Spannung zusätzliche Normalspannungen wirksam werden.

6.4.8.2. Diffusion. In Kristallen erfolgt sie nach speziellen Platzwechselmechanismen. Platzwechselvorgänge können in aus identischen Atomen aufgebauten Kristallen durch Ausgleich verschiedener kinetischer Energien der Atome zustandekommen. Man spricht hier von *Selbstdiffusion.* Damit ein Gitterbaustein seinen normalen Platz verlassen kann, muß eine bestimmte Aktivierungsenergie aufgebracht werden. Sie ist besonders niedrig für an der Oberfläche befindliche Kristallbausteine, da diese eine geringere Bindungsenergie besitzen als die Atome im Kristallinneren. Die Aktivierungsenergien für Platzwechselvorgänge im Kristallinneren werden durch die Realstruktur entscheidend beeinflußt. Je nach dem Elementarprozeß der Teilchenverschiebung unterscheidet man verschiedene Arten. Bei der *Diffusion über Leerstellen* (Abb. 6.4.8-7a) erfolgt der Austausch zwischen Gitterbaustein und Leerstelle. Energetisch wie geometrisch ist dieser Mechanismus bevorzugt, weil als Aktivierungsenergie nur die Ablösearbeit der Gitterbausteine aufgebracht werden muß. Leerstellen sind in einem Kristall bei Temperaturen oberhalb 0 K immer vorhanden (↑ 6.4.7.3.1.). Bei der *Diffusion über Zwischengitteratome* (Abb. 6.4.8-7b) wandern die Atome auf Zwischengitterplätzen. Die Bewegung erfordert eine starke Gitterverzerrung, die mit einer hohen Aktivierungsenergie verbunden ist. Die *Diffusion durch Ringsprünge* (Abb. 6.4.8-7c) erfolgt durch einen Ringtauschmechanismus, an dem mehr als zwei Atome beteiligt sind. Er wurde von ZENER vorgeschlagen. Bei der *Diffusion durch direkten Austausch* (Abb. 6.4.8-7d) ist der direkte Platztausch gegenüber dem Ringtausch mit einer so starken Gitterverzerrung verbunden, daß die hohe Aktivierungsenergie durch thermische Energie allein kaum aufgebracht werden kann. Solche Platzwechselvorgänge sind daher unwahrscheinlich.

Die Aktivierungsenergien für verschiedene Mechanismen der Diffusion im kubisch-flächenzentrierten Gitter des Kupfers (Tab. 6.4.8-8) zeigen, daß die Leerstellendiffusion begünstigt ist. Auch die experimentell erhaltene Aktivierungsenergie kommt dem dafür geltenden Wert nahe.



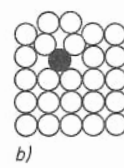






Abb. 6.4.8-7 Platzwechselmechanismen in Kristallen aus einer Atomsorte

Tab. 6.4.8-8 Diffusionsmechanismen in Kupfer und zugehörige berechnete Aktivierungsenergien E_A

Diffusionstyp	E_A in kJ mol^{-1}
Leerstellendiffusion	270
Ringdiffusion	380
Zwischengitterdiffusion	960
direkter Platztausch	1060

Im Bereich von Korngrenzen ist die Diffusion erleichtert gegenüber der in einem nur mit Punktdefekten behafteten Kristall. Die Aktivierungsenergie der Selbstdiffusion kann bei Metallen wie Blei auf die Hälfte, bei Nickel sogar auf ein Viertel absinken. Da mit abnehmender Aktivierungsenergie die Diffusionskoeffizienten D zunehmen, ergibt sich im allgemeinen $D_{(\mathrm{Oberfläche})} > D_{(\mathrm{Korngrenze})} > D_{(\mathrm{Gitter})}$.

koeffizienten in den drei kristallographischen Achsen mit α_a, α_b und α_c, ergibt sich folgende Zuordnung:

Kubisches System: $\alpha_a = \alpha_b = \alpha_c$; die thermische Ausdehnung ist isotrop, das Ellipsoid entartet zur Kugel.

Hexa-, tri- und tetragonales System: $\alpha_a = \alpha_b \neq \alpha_c$; d. h., die Ausdehnung wird durch ein Rotationsellipsoid beschrieben.

Tab. 6.4.8-9 Lineare thermische Ausdehnungskoeffizienten α in K^{-1} einiger Kristallarten (0 bis 100 °C); $\alpha_{\parallel}$ Ausdehnungskoeffizient $\parallel$ c-Achse, $\alpha_{\perp}$ Ausdehnungskoeffizient $\perp$ c-Achse

Mineral	Kristallklasse			
		α		
Steinsalz (NaCl)	$m3m$	$40 \cdot 10^{-6}$		
Fluorit (CaF_2)	$m3m$	$19 \cdot 10^{-6}$		
		$\alpha_{\parallel}$	$\alpha_{\perp}$	
Cd	$6/mmm$	$49 \cdot 10^{-6}$	$17 \cdot 10^{-6}$	
Zn	$6/mmm$	$55 \cdot 10^{-6}$	$14 \cdot 10^{-6}$	
Calcit ($CaCO_3$)	$3m$	$26 \cdot 10^{-6}$	$-6 \cdot 10^{-6}$	
Quarz (SiO_2)	32	$9 \cdot 10^{-6}$	$14 \cdot 10^{-6}$	
		α_a	α_b	α_c
Aragonit ($CaCO_3$)	mmm	$10 \cdot 10^{-6}$	$16 \cdot 10^{-6}$	$33 \cdot 10^{-6}$
Chrysoberyll (Al_2BeO_4)	mmm	$6{,}0 \cdot 10^{-6}$	$6{,}0 \cdot 10^{-6}$	$5{,}2 \cdot 10^{-6}$

6.4.8.3. Thermische Ausdehnung. Die lineare thermische Ausdehnung ist eine besonders charakteristische anisotrope Kristalleigenschaft. Erwärmt man einen aus einem Kristall herausgeschnittenen Stab der Länge l_0 um die Temperaturdifferenz $\Delta T = T - T_0$, dann gilt für die neue Länge l des Stabs bei T die Beziehung $l = l_0(1 + \alpha \Delta T)$. Dabei wird α als *linearer thermischer Ausdehnungskoeffizient* bezeichnet. Tab. 6.4.8-9 zeigt, daß bei den nichtkubischen Kristallen der Ausdehnungskoeffizient teilweise stark richtungsabhängig ist. Wird eine aus einem Einkristall herausgeschliffene Kugel erwärmt, bleibt wegen der Anisotropie von α ihre Gestalt im allgemeinen nicht erhalten; sie wird zu einem Ellipsoid. Es besteht ein direkter Zusammenhang zwischen der Symmetrie der Ausdehnungskörper und der der Kristalle. Bezeichnet man die linearen thermischen Ausdehnungs-

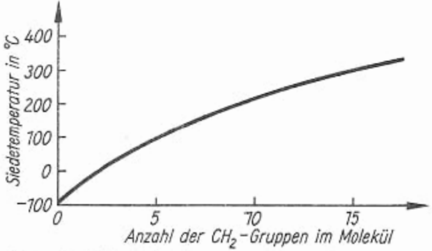


Abb. 6.5.1-1 Siedetemperaturen der unverzweigten Paraffinkohlenwasserstoffe (Alkane) als Funktion der Kettenlänge

Rhombisches, mono- und triklines System: $\alpha_a \neq \alpha_b \neq \alpha_c$; der Ausdehnungskörper ist ein dreiachsiges Ellipsoid.

6.5. Zustandsformen makromolekularer Stoffe

6.5.1. Besonderheiten makromolekularer Stoffe

6.5.1.1. Kohäsionsenergie und Molekülgröße. Je größer die in einem Stoff wirkenden zwischenmolekularen Kräfte sind, um so mehr Energie ist erforderlich, um Teilchen voneinander zu trennen. Die *Kohäsionskräfte* zwischen Makromolekülen sind qualitativ die gleichen wie zwischen einfachen Molekülen, im Polyethylen also z. B. dieselben wie in niedermolekularen Paraffinen. Während aber bei diesen nur wenige CH_2- und CH_3-Gruppen in jedem Molekül vorhanden und für die Wechselwirkungen verfügbar sind, beruht die Kohäsion zwischen Polyethylenmolekülen auf Tausenden von CH_2-Gruppen. Mit der Zahl der CH_2-Gruppen steigt die Siedetemperatur (Abb. 6.5.1-1). Danach läßt sich für Polyethylen durch Verlängerung der Kurve eine sehr hohe Siedetemperatur abschätzen, die sich experimentell jedoch nicht erreichen läßt, weil das Material schon vorher thermisch abgebaut wird. Makromolekulare Stoffe existieren daher nicht in der Gasphase.

6.5.1.2. Platzwechselvorgänge. Fließ- und Diffusionsvorgänge in kompakten Materialien erfordern auch in Systemen, in denen nur eine Nahordnung vorliegt, Übergänge der Teilchen

von einem Gitterpunkt zu einem anderen. In niedermolekularen Stoffen mit Molekülgittern tauschen dabei jeweils *Moleküle* ihre Positionen. In hochmolekularen Stoffen sind die Gitterplätze mit *Kettengliedern* besetzt. Da jedes Glied mit zwei Nachbarn verknüpft ist, tritt an Stelle der freien Diffusion die *gebundene Diffusion*, bei der ein Platzwechsel nur unter Mitführung der Nachbarglieder möglich ist. Diese Kopplung erschwert die Austauschvorgänge erheblich, was z. B. die sehr hohe Viskosität makromolekularer Schmelzen zur Folge hat. Die Beweglichkeit kann so gering werden, daß sich das polymere Material bei Temperaturen, bei denen das Monomer eine leicht bewegliche Flüssigkeit ist, wie ein fester Stoff verhält.

6.5.1.3. Grundsätzliche Unterschiede zwischen vernetzten und unvernetzten Polymeren. Bei niedrigen Temperaturen hat ein vernetzter Plastwerkstoff in vieler Hinsicht gleiche Eigenschaften wie ein Kunststoff, der aus Kettenmolekülen besteht. Anders in der Wärme: Wenn die Molekularbewegungen lebhaft genug sind, können die Moleküle im unvernetzten Material ihre gegenseitige Lage verändern. Linearpolymere werden daher beim Erwärmen prinzipiell plastisch, sie sind *Thermoplaste*. Bei einer vernetzten Struktur betrifft die Erleichterung der Platzwechselvorgänge dagegen nur die Molekülteile zwischen jeweils zwei Vernetzungsstellen. Gleitvorgänge sind durch das räumliche Netzwerk unmöglich, daher werden vernetzte Polymere nicht plastisch. Wenn die Maschen des Netzes eng sind, bleiben die Stoffe bis zu ihrem Zersetzungspunkt hart und sind *Duroplaste*. Sind die Maschen dagegen sehr weit, dann werden auch vernetzte Polymere bei hinreichend hoher Temperatur weich. Durch äußere Kräfte kann das Netzwerk verzerrt werden. Beim Wegfall der Krafteinwirkung nehmen die Kettenteile ihre vorherige Konformation (↑ 5.4.6.2.) und der Körper seine Ausgangsgestalt wieder an. Er verhält sich elastisch, das Material gehört zur Gruppe der *Elaste* oder *Elastomeren*. Wird zu hoch erhitzt und das Netzwerk durch thermische und mechanische Einflüsse zerstört, dann ist das Material irreversibel geschädigt.

Eine interessante und technisch wichtige Sonderstellung nehmen *Ionomere* (↑ 5.4.2.1.) ein, in denen die Makromoleküle über Ionenbeziehungen vernetzt sind. Wenn diese durch thermisch-mechanische Einwirkungen aufgebrochen werden, so können die im übrigen intakten Ketten gegeneinander gleiten. Das Material kann spanlos umgeformt werden. Beim Abkühlen bilden sich die Ionenbeziehungen zwischen anderen Ladungsträgern erneut aus: Das Material ist wieder vernetzt. Ähnlich wirkt *physikalische Vernetzung* durch kristalline Bereiche.

Thermoplaste können dank ihrer Fließfähigkeit in der Wärme durch *Spritzgießen* verarbeitet werden. Dieser Vorgang kann im Prinzip beliebig wiederholt werden, z. B. bei der Verwertung von Ausschußstücken und Abfällen. Duroplaste und Elaste werden als *Zwischenprodukte* ausgeformt und erst danach *vernetzt*. Dieser Verarbeitungsprozeß umfaßt neben physikalischen auch chemische Vorgänge. Die Vernetzung wird bei den Duroplasten *Aushärtung* und bei den Elasten *Vulkanisation* genannt.

Linearpolymere können sich in geeigneten Flüssigkeiten *rückstandslos auflösen*. Vernetzte Polymere sind dagegen grundsätzlich nur quellbar, wobei der maximal erreichbare *Quellungsgrad* mit der Maschenweite zunimmt.

6.5.2. Einfluß der Temperatur auf den Zustand der Thermoplaste

Die Eigenschaften unvernetzter Plastwerkstoffe hängen stark von der Temperatur ab und durchlaufen beim Übergang vom Sprödverhalten in der Kälte bis zum plastischen Fließen in der Wärme mehrere Zustandsbereiche. Die Breite der Bereiche ist von Material zu Material verschieden: bei Polyamiden z. B. ist das Intervall für das zähfeste Verhalten sehr breit und ein elastischer Bereich praktisch gar nicht vorhanden; beim Polyisobutylen ist das Gebiet der Plastizität breit, beim Polyvinylchlorid dagegen eng und beim Polyacrylnitril nicht auffindbar, da es von der beginnenden Zersetzung überdeckt wird. Der Übergang von einem Zustand zum benachbarten kann an der Änderung von makroskopischen Eigenschaften erkannt werden, die aber in der Regel unterschiedliche Abhängigkeit von der molekularen Beweglichkeit, von den zwischenmolekularen Kräften usw. aufweisen. Verschiedene Meßverfahren können daher verschiedene Werte für die Umwandlungstemperaturen ergeben.

Die Anwendung von Thermoplastformkörpern wird durch das *Erweichen* bei hohen Temperaturen begrenzt. Die Grenztemperatur für die *Formbeständigkeit in der Wärme* liegt mehr oder weniger knapp unterhalb des Plastizitätsgebiets.

Die Formgebung erfolgt im plastischen Zustand, wobei die höchsten zulässigen Temperaturen durch den *Zersetzungsbeginn* gegeben sind. Nach tiefen Temperaturen hin wird das Einsatzgebiet des Materials häufig durch eine *Versprödung* eingeengt.

6.5.3. Glaszustand und Einfriertemperatur

Als Kennzeichen des Glaszustands liegt die *Sprödigkeit* nahe, die vom Fensterglas her vertraut ist und tatsächlich auch bei manchen Poly-

meren und anderen glasig erstarrenden Materialien auftritt. Doch gibt es Polymere, die glasklar und dabei so zähfest sind, daß sie für die Gesichtsfelder der Kosmonautenhelme verwendet werden können, die einen Meteoritenaufschlag bei Weltraumtemperaturen überstehen müssen.

6.5.3.1. Einfriertemperatur. Zur Abgrenzung des Glaszustands ist die Untersuchung des Einfriervorgangs nützlich.

Beim Abkühlen der Schmelze eines glasig erstarrenden Stoffs nimmt das Volumen etwa proportional mit der Temperatur ab (Abb. 6.5.3-1a). Bei tiefen Temperaturen ist die Volumenkontraktion ebenfalls linear, aber deutlich geringer. Die beiden geraden Kurvenäste schneiden sich in einem Punkt, dessen Abszissenwert die *Einfriertemperatur* T_E ergibt, unterhalb der sich der Stoff im *Glaszustand* befindet. Die Kurve für die Enthalpie H verläuft ganz ähnlich.

Die Temperaturkoeffizienten der Meßgrößen V und H, der *Ausdehnungskoeffizient* $\mathrm{d}V/V\,\mathrm{d}T = \alpha$ bzw. die spezifische Wärme $\mathrm{d}H/\mathrm{d}T = c$ weisen bei der Einfriertemperatur eine sprunghafte Änderung auf (Abb. 6.5.3-1b). Der Übergang vom steilen zum flachen Ast der V-T-Kurven tritt bei tieferen Temperaturen auf, wenn das Volumen erst nach längeren Wartezeiten nach jeder Temperaturänderung gemessen wird. Der in Abb. 6.5.3-1a dargestellte Verlauf kann daher nicht auf thermodynamischen Gründen beruhen; er geht vielmehr auf ein *kinetisches Phänomen* zurück: Oberhalb T_E nimmt das Volumen bei jeder Temperatur den Gleichgewichtswert an. Die Gleichgewichtslagen für tiefere Temperaturen würden auf einer Verlängerung des steilen Kurvenasts liegen. Das Volumen eines glasig erstarrten Polymers ist demnach größer als das *Gleichgewichtsvolumen* (Abb. 6.5.3-2). Wenn eine Probe von T_1 auf T_2 abgekühlt wird, so ist V_2 das neue Gleichgewichtsvolumen, das in zwei Etappen erreicht wird: Mit der Temperaturerniedrigung tritt eine *momentane Kontraktion* von V_1 auf V_1' ein. Für $T_2 < T_E$ liegt V_1' auf dem sofort meßbaren Kurvenast des glasig erstarrten Materials. Der Übergang von V_1' zu V_2 erfolgt um so langsamer, je tiefer T_2 unter T_E liegt. Es kann Jahre dauern, bis der *Ausgleichsvorgang* beendet ist. Bei der Aufnahme einer Volumen-Temperatur-Kurve wird im allgemeinen mit Wartezeiten von einigen Minuten gearbeitet, so daß nur Ausgleichsvorgänge mit relativ hoher Geschwindigkeit erfaßt werden. Unterhalb einer bestimmten Grenztemperatur läuft der Vorgang zu langsam ab, und man findet die auf dem flachen Kurvenast liegenden Volumina V_1'.

6.5.3.2. Molekulare Deutung des Einfrierphänomens. Die Teile der Makromoleküle schwingen um ihre Ruhelage mit einer Amplitude, die mit sinkender Temperatur abnimmt. Jedes Glied hat daher bei tieferen Temperaturen einen geringeren Platzbedarf als bei höheren. Je kleiner das von jedem Bauelement beanspruchte Volumen ist, um so enger können die Ketten zusammenrücken (Abb. 6.5.3-3). Diese Umlagerung, die *Mikro-Brownsche-Bewegung*, braucht im Unterschied zur Makro-Brownschen Bewegung, die z. B. beim Fließen auftritt, eine gewisse Zeit, die bei hoher Temperatur kurz und bei tiefer lang ist.

Die Mikro-Brownsche Beweglichkeit wird durch Überladung der Kette mit Substituenten behindert, daher liegt T_E für Polymethylmethacrylat über 60 K höher als für Polymethylacrylat. Analog wirken starke zwischenmolekulare Kräfte wie z. B. Wasserstoffbrücken;

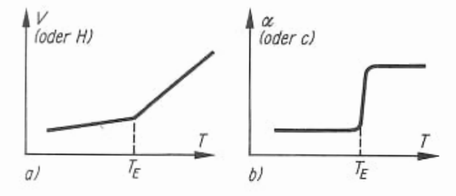


Abb. 6.5.3-1 *a*) Volumen V oder Wärmeinhalt H eines glasartig erstarrenden Materials als Funktion der Temperatur T (schematisch), *b*) Änderung des Ausdehnungskoeffizienten α bzw. der spezifischen Wärme c mit der Temperatur T für glasig erstarrende Polymere; T_E Einfriertemperatur

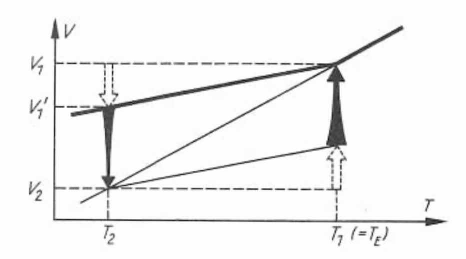


Abb. 6.5.3-2 Volumenrelaxation: Änderung des Volumens V mit der Temperatur T; schematische Aufteilung in momentane Änderung und Nachwirkung, die *Pfeildicke* deutet die Geschwindigkeit an

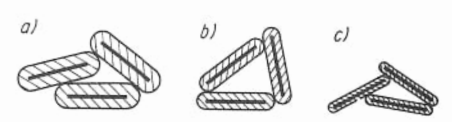


Abb. 6.5.3-3 Änderung des Volumens mit der Temperatur (schematisch); a) oberhalb, *b*) bei, *c*) unterhalb der Einfriertemperatur

Polyvinylalkohol hat eine um 55 K höhere Einfriertemperatur als Polyvinylacetat.

6.5.3.3. Dämpfung von Schwingungen. Wirkt auf eine polare Substanz ein elektrisches Wechselfeld, so orientieren sich die Segmente in Feldrichtung. Ist die Temperatur und damit die molekulare Beweglichkeit genügend hoch, so können sich die Segmente im Takt mit dem angelegten Feld in die jeweils günstigste Lage einstellen, und es wird praktisch keine Energie absorbiert. Ist andererseits die Temperatur niedrig und die Beweglichkeit entsprechend klein, können sich die Segmente nicht mehr bewegen, so daß ebenfalls nur wenig Energie aufgenommen wird. Zwischen diesen beiden Gebieten liegt ein Temperaturbereich, in dem die Segmente dem äußeren Wechselfeld mit einer Phasenverschiebung folgen. Jetzt wird dem angelegten Feld Energie entnommen, das Material hat in diesem Zustand einen hohen Wert des *dielektrischen Verlustfaktors* tan δ. Das Maximum der Energieabsorption liegt knapp oberhalb der Einfriertemperatur. Mit steigender Frequenz wird zunehmend höhere Beweglichkeit für verlustfreies Nachstellen der Segmente erforderlich. Dadurch verschiebt sich das Maximum nach höheren Temperaturen (Abb. 6.5.3-4). Auch bei erzwungenen mechanischen

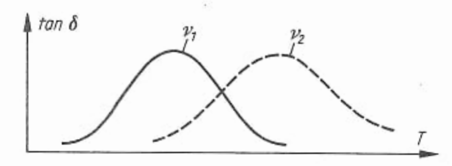


Abb. 6.5.3-4 Abhängigkeit des dielektrischen Verlustfaktors tan δ von der Temperatur T (schematisch) für zwei Frequenzen ($\nu_1 < \nu_2$)

Schwingungen tritt eine temperaturabhängige Energieabsorption mit einem Maximum in der Nähe der Einfriertemperatur auf. Trägt man die Temperatur des Maximums über dem Kehrwert der Anregungsfrequenz, d. h. dem Zeitintervall für jeweils eine Schwingung auf, so ergeben die bei Sekundenbruchteilen liegenden Punkte dieser Messungen Kurven, die sich erstaunlich gut zur statisch bestimmten Einfriertemperatur mit beispielsweise 10 min Nachstellzeit verlängern lassen.

6.5.4. Kautschukelastizität

Kautschukelastische Körper werden schon durch verhältnismäßig kleine Kräfte stark deformiert. Wenn die Kräfteeinwirkung aufhört, nehmen sie wieder ihre ursprüngliche Form an, sie besitzen *Gestaltelastizität*. Die für eine bestimmte elastische Formänderung erforderliche Belastung wird durch den *Elastizitätsmodul* gekennzeichnet, der für gummielastische Materialien viel kleiner ist als für andere elastische Körper; der für Kautschuk ist z. B. der millionste Teil des Werts von Stahl. Ein weiteres Merkmal ist die *große reversible Dehnbarkeit*. Kautschuk kann um etwa 800% seiner ursprünglichen Länge gedehnt werden, ohne bleibende Formänderungen zu erleiden oder zu reißen. Beim gleichfalls hochelastischen Stahl beträgt die elastische Dehnung etwa 0,1% (s.a. 6.4.8.1.)

6.5.4.1. Struktur kautschukelastischer Materialien. Kautschukelastizität ist eine *Konformationselastizität* oder, im Unterschied zur Energieelastizität der Metalle, eine *Entropieelastizität*. Sie tritt nur oberhalb der Einfriertemperatur bei hoher Segmentbeweglichkeit auf (↑ 6.5.3.2.) und ist ursächlich mit der Knäuelkonformation in amorphen Polymeren verbunden (↑ 5.4.6.2.2.). Veränderung der Gestalt durch äußere Kräfte ruft eine Triebkraft zur Rückkehr in die Knäuelkonformation hervor.

Gummielastische Stoffe sind vernetzt. Besonders gute Kautschukeigenschaften zeigen schwach vernetzte Polymere mit niedrigen Einfriertemperaturen. Hierzu gehören vulkanisierter Naturkautschuk und die synthetisch erzeugten Kautschuktypen im verarbeiteten Zustand. Auch bei diesen Stoffen ist das gummiartige Verhalten an ein bestimmtes Temperaturintervall geknüpft. Naturkautschuk verliert z. B. seine Elastizität bei −50 °C. Auch Thermoplaste können zwischen Einfriertemperatur und Erweichungsbeginn gummiartige Eigenschaften zeigen. Bei geeigneter Lage dieser Temperaturen sind die Materialien im normalen Anwendungsgebiet kautschukelastisch, z. B. Weich-PVC oder Ethylen-Vinylacetat-Copolymere. Bei diesen Stoffen verhindern zwischenmolekulare Kräfte das Abgleiten der Ketten bei Dehnung, es liegt *physikalische Vernetzung* vor. Die Haftpunkte sichern die makroskopische Gestalttreue; die Knäuelungstendenz der Kettenstücke kann sich zur meßbaren Rückstellkraft summieren, die den gedehnten Körper in seine ursprüngliche Form zurückschnellen läßt, sobald die äußere Einwirkung wegfällt (Abb. 6.5.4-1).

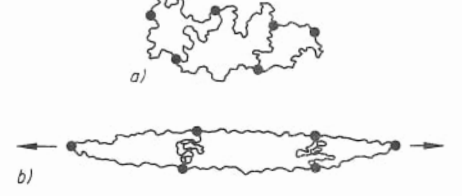


Abb. 6.5.4-1 Entropieelastizität eines makromolekularen Netzwerks; *a*) ungedehnt, *b*) gedehnt

Maschenweite des Netzwerks und Beweglichkeit der Netzbögen bestimmen den Elastizitätsmodul des Materials. Neuere Erkenntnisse über diese Zusammenhänge haben zur Entwicklung von *Hartelastomeren* geführt, die die Lücke zwischen Stahl und Gummi überbrücken und für kraftübertragende Dämpfungselemente gebraucht werden.

6.5.4.2. Eingefrorene Deformationen. Wird Gummi im gedehnten Zustand unter die Einfriertemperatur abgekühlt, so wird den Molekülen die Möglichkeit zur Rückkehr in die Knäuelformation genommen. Die Deformation bleibt auch nach Aufhören der Krafteinwirkung bestehen. Wird die Segmentbeweglichkeit durch Wärmezufuhr erhöht, dann wird das Material wieder elastisch und nimmt seine frühere Gestalt an. Es hat den Anschein, als ob es sich plötzlich auf seine ursprüngliche Form besinnen würde; man spricht vom *Erinnerungseffekt*. Dieses Prinzip wird bei Schrumpfkörpern und Schrumpfverpackungen genutzt, die mit eingefrorenen Dehnungen in den Handel gebracht werden und vom Verbraucher durch Erwärmen auf einen bestimmten Gegenstand aufgeschrumpft werden können.

6.5.5. Schmelzen und Kristallisieren

6.5.5.1. Geordnete Bereiche. Die Feinstrukturuntersuchung von Polymeren mit Röntgenstrahlen läßt bei einigen Stoffen erkennen, daß neben einer ungeordneten, *amorphen* Grundstruktur auch geordnete, *kristalline* Bereiche vorliegen. Wenn die Ordnung auch weit hinter der eines Idealkristalls zurückbleibt, so ist doch eine weitgehende Regelmäßigkeit der Lage der Strukturelemente in drei, nicht in einer Ebene liegenden Raumrichtungen unverkennbar.

Nach den Röntgenuntersuchungen liegt die Größe der geordneten Bereiche allerdings wesentlich unter der Länge eines gestreckten Makromoleküls. Dieser überraschende Befund hat zur Entwicklung verschiedener Modellvorstellungen Anlaß gegeben. Im *Fransenmodell* wird angenommen, daß in den geordneten Bereichen jeweils nur verhältnismäßig kurze Teile von verschiedenen Makromolekülen vereinigt sind (Abb. 6.5.5-1a). Jede Kette gehört dann mehreren verschiedenen kristallinen Bereichen an, zwischen denen amorphe Bezirke liegen. Das *Lamellenmodell* wurde an Hand von Erkenntnissen entworfen, die an Einkristallen gewonnen worden sind. Es enthält die Vorstellung, daß eine Kette ein und denselben Kristallit mehrmals durchläuft. Da die Umkehrstellen nicht ideal kurz ausgebildet sind, sondern unregelmäßige Schlaufen darstellen, ist das kristalline Innere der Lamellen beiderseits in dünne,

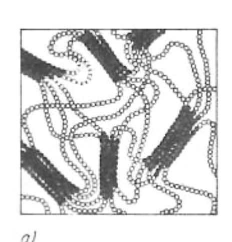




Abb. 6.5.5-1 *a)* Fransenmodell, *b)* Ausschnitt aus einer Lamelle mit ungleichförmig gefalteten Ketten; Kettenteile in ungeordneten Bereichen *offen*, Kettenteile in geordneten Bereichen *schwarz*

ungeordnete Schichten eingebettet (↑ Abb. 6.5.5-1b).

6.5.5.2. Einkristalle. Bei langsamer Abscheidung von Polyethylen aus sehr verdünnten Lösungen entstehen regelmäßig geformte Niederschläge von hohem Kristallisationsgrad. Das Elektronenmikroskop zeigt ihren Aufbau aus dünnen, rautenförmigen Blättchen, die terassenförmig übereinanderliegen können (Tafeln 30/31). Die Dicke der Lamellen hängt von der Kristallisationstemperatur ab und liegt bei etwa 10 nm. In den Lamellen laufen die Ketten im rechten Winkel zu den Deckflächen. Ähnliche Einkristalle lassen sich unter geeigneten Bedingungen aus allen kristallisationsfähigen Polymeren züchten.

6.5.5.3. Kristallitschmelzpunkt. Die Volumenzunahme von kristallisierten Polymeren wird mit steigender Temperatur immer größer, bis die *V-T*-Kurve schließlich beim Kristallitschmelzpunkt einen Knick aufweist und in eine Gerade mit wesentlich geringerer Steigung übergeht. In dem etwa 100 K umfassenden *Schmelzintervall*, das mit dem Ansteigen der *V-T*-Kurve beginnt und beim Knickpunkt endet, werden die geordneten Bereiche kleiner und kleiner, ihre Struktur ist zunehmend gestört. Oberhalb des Knickpunkts sind im Material keine geordneten Bereiche mehr nachweisbar.

Der *Ordnungsgrad*, der Anteil an räumlich geordnetem Material, kann nach mehreren Methoden bestimmt werden: röntgenographisch, aus der Intensität bestimmter Ultrarotabsorptionsbanden, durch Kernresonanzmessungen, durch kalorimetrische Untersuchungen im Schmelzbereich oder einfach durch Dichtemessungen. Die Ergebnisse weichen z. T. beträchtlich voneinander ab.

Die Ausbildung der geordneten Bereiche ist an bestimmte strukturelle Voraussetzungen gebunden (↑ 5.4.6.2.1.) und erfordert Zeit. Die *Kristallisationsgeschwindigkeit* wächst mit fallender Temperatur vom Wert Null im Kristallitschmelzpunkt rasch bis zu einem Maximalwert

und geht bei tieferen Temperaturen wieder bis auf unmeßbar kleine Werte zurück. Daher kann durch Abschrecken ein ungeordneter Zustand eingefroren werden. Ein abgeschrecktes Polymer kristallisiert bei Wärme- oder Quellmitteleinwirkung und erhält dabei andere Eigenschaften.

6.5.5.4. Überstrukturen. Partiell kristalline Hochpolymere sind unterhalb des Kristallitschmelzpunktes *opak*, d. h. undurchsichtig, während ihre Schmelzen wasserklar sind. Die Trübung, die bei der Abkühlung auftritt, hängt offensichtlich mit der Kristallisation zusammen. Andererseits ist durch röntgenographische Messungen bekannt, daß die geordneten Bereiche wesentlich kleiner als die Wellenlängen des sichtbaren Lichts sind und daher die Trübung nicht unmittelbar hervorrufen können. In kristallinen Polymeren müssen folglich noch andere Inhomogenitäten der optischen Dichte vorliegen, die größer als die Kristallite sind. Die geordneten Bereiche können sich zu kugelförmigen Gebilden, den *Sphärolithen*, zusammenlagern. Diese bestehen aus einer großen Zahl von unregelmäßigen Strahlen, die sternförmig von einem Keim ausgehen. Je mehr Keime in einer Schmelze vorliegen, desto *mehr* und desto *kleinere* Sphärolithe entstehen. In durchkristallisierten Materialien sind die Sphärolithe zusammengewachsen und haben Polyedergestalt. Im Polarisationsmikroskop sind ihre Umrisse und die periodische Orientierung im Innern zu erkennen.

Die mechanischen Eigenschaften werden durch Größe und Anzahl der Sphärolithe beeinflußt. Wenige große bewirken Härte, aber auch Versprödung, viele kleine Zähfestigkeit. Diese Erkenntnisse werden genutzt, um bestimmte Werkstoffeigenschaften durch *gesteuerte Kristallisation* zu erzielen.

In Fasern treten langgestreckte Überstrukturen auf, die *Fibrillen*, die ebenfalls aus vielen Kristalliten bestehen.

6.5.6. Orientierte Polymere

Die *Kraft-Dehnungs-Diagramme* von makromolekularen Werkstoffen lassen häufig einen *Verstreckungseffekt* erkennen. Sie stellen den Zusammenhang zwischen der langsam ansteigenden Belastung beim Zugversuch und der dadurch hervorgerufenen Dehnung graphisch dar. Bei verstreckbaren Polymeren folgt auf eine steile Zunahme der Belastung bei verhältnismäßig geringen Dehnungen eine *Fließzone*, in der sich die Probe bei gleichbleibender Belastung um ein Mehrfaches verlängert. Durch die Verstreckung verringert sich der Querschnitt, während die Festigkeit erstaunlicherweise größer wird. Die Festigkeitszunahme führt beim Verstrecken eines Formkörpers zur Ausbildung einer Schulter. Hier geht das noch unverstreckte Material in den schmalen Strang des verstreckten über, der entsprechend der Querschnittsverminderung mit 3- bis 5fach höherer Geschwindigkeit als die Schulter läuft. Dieser Vorgang erinnert an das Ausziehen eines alten Fernrohrs und wird *Teleskopeffekt* genannt. Die erzielbare Festigkeitssteigerung wird technisch bei der Produktion von Fasern und Folien, in einigen Fällen auch bei der Herstellung von Rohren und Platten genutzt.

6.5.6.1. Molekulare Deutung des Reckvorgangs. Die Struktur eines unverstreckten Thermoplasten mit Knäuelkonformation der einzelnen Moleküle kann mit einem Wattebausch verglichen werden. Wenn aus diesem ein Faden gezupft wird, so lagern sich darin die Einzelfasern mehr oder minder parallel. Das gleiche erfolgt beim Streckprozeß mit den Makromolekülen. Im polarisierten Licht ist die Orientierung zu erkennen. Wird ein orientierter Körper in Streckrichtung auf Zug beansprucht, so wird die Kraft vornehmlich von den Hauptvalenzen aufgenommen, die um eine Zehnerpotenz stärker sind als die Nebenvalenzen. Daraus resultiert die Festigkeitszunahme trotz der Querschnittsverminderung.

6.5.7. Eingefrorene Spannungen und Rückstellvorgänge

Bei der Spritzgußverarbeitung wird der zähflüssige Thermoplast durch hohen Druck in die Form gepreßt. Bei langen Fließwegen und schmalen Fließkanälen orientieren sich die Ketten dabei in Fließrichtung. Erstarrt die Masse rasch, so bleibt den Molekülen keine Zeit, in die Gleichgewichtskonformation zurückzukehren. Die geringe Geschwindigkeit der Molekularbewegungen verhindert den Abbau der inneren Spannungen. Diese können unbemerkt bleiben, bis irgendwann äußere Einflüsse wie Temperaturerhöhung oder das Eindringen von Quellmitteln die Beweglichkeit der Ketten erhöhen. Dann gleichen sich die inneren Spannungen aus, was äußerlich als *Verziehen* bei Spritzgußteilen oder als *Einlaufen* bei Textilien (Wärme und Quellmitteleinwirkung beim Waschen) sichtbar wird. Verzögerte Molekularbewegungen, die zu einem allmählichen Abbau von eingefrorenen Spannungen führen, lösen oft Rißbildung parallel zur Orientierungsrichtung aus, besonders wenn durch Wärmeausdehnung oder Anquellen noch zusätzliche Spannungen geschaffen werden. In Fließrichtung verlaufende Alterungsrisse finden sich häufig an Polystyrengefäßen, die mit heißen oder sehr kalten Medien in Berührung gekommen sind. Die leichte *Spaltbarkeit orientierter Folien*

wird technisch zur Faserstoffherstellung genutzt.
Plötzliche Rißbildung in einem Werkstoff, auf den mechanische Spannungen und chemische Einflüsse gleichzeitig einwirken, bezeichnet man als *Spannungsrißkorrosion* (Tafeln 42/43). Zum Unterschied zu der eben erwähnten Bildung von Quellungsrissen können in diesem Fall auch nichtquellende Chemikalien eine große Wirkung entfalten. Auf Polyethylen haben z. B. Netzmittellösungen starken Einfluß.
Die inneren Spannungen in verstreckten Polymeren lassen sich durch Erwärmen unter Belastung abbauen, ohne daß Schrumpf erfolgt oder die Orientierung verloren geht. Bei diesem *Fixiervorgang* spielen Ordnungsprozesse eine Rolle. Ein fixiertes Material widersteht im Idealfall Wärmeeinwirkungen bis zur Höhe der Fixiertemperatur, ohne einzulaufen.

6.5.8. Hochpolymere Schmelzen

Bei Temperaturen um 200 °C bilden die Thermoplaste zähflüssige Schmelzen. Da sich solche Temperaturen ohne Schwierigkeiten beherrschen lassen, werden diese Materialien überwiegend im plastischen Zustand verarbeitet.
Die Zähigkeit η der hochpolymeren Schmelzen nimmt, wie die aller Flüssigkeiten, mit steigender Temperatur T exponentiell ab. Häufig gilt die Gleichung $\eta = A \exp [B/(T - C)]$ mit den Konstanten A, B, C. Mit steigender *molarer Masse* $\bar{M}_w$ (↑ 5.4.3.3.) nimmt die Zähigkeit der Schmelzen stark zu. Mit den Konstanten K und ε gilt in vielen Fällen $\lg \eta = \lg K + \varepsilon \lg \bar{M}_w$. ε liegt bei 3,1 bis 3,4.

6.5.8.1. Existenzbereich der Schmelzen. Niedermolekulare Flüssigkeiten erstarren bei einer ganz bestimmten Temperatur. Im Schmelzgleichgewicht ist das harte, kristalline Material mit Leichtigkeit neben der beweglichen Flüssigkeit zu erkennen. Eine hochpolymere Schmelze hingegen wird beim Abkühlen *allmählich* zäher, bis sie schließlich wie ein fester Stoff erscheint. Der Verfestigungsprozeß überstreicht ein Intervall von 100 K oder mehr; an seinem oberen Ende kann das Material im *Spritzguß* verarbeitet werden. Im Mittelbereich ist *thermoplastische Umformung* durch Biegen oder Ziehen möglich, und am unteren Ende ist die Formbeständigkeit erreicht.
Kristallisierende Polymere mit hohen Kristallitschmelzpunkten und molaren Massen um 10000 g mol^{-1}, z. B. Polyamide, werden aus relativ dünnflüssigen Schmelzen heraus in einem wesentlich engeren Intervall fest. Die obere Begrenzung des Existenzgebiets der Schmelzen ist durch die Zersetzung bestimmt. Manche Polymere sind bei Ausschluß von Sauerstoff bis zu hohen Temperaturen stabil.

6.5.9. Hochpolymere im gequollenen Zustand und in Lösung

6.5.9.1. Verhalten gegen Lösungsmittel. *Similia similibus solventur* – ein Stoff wird durch ihm ähnliche gelöst – diese jahrhundertealte Regel der Chemiker gilt auch für makromolekulare Verbindungen. In *Wasser* sind nur die Polymere löslich, die reichlich wasserfreundliche Gruppen, z. B. —OH, —COOH, besitzen; andere Polymere lösen sich in *organischen Lösungsmitteln*, die ihnen strukturell und in physikalisch-chemischen Eigenschaften ähneln. Manche Linearpolymere, wie Polytetrafluorethylen, werden durch so starke zwischenmolekulare Kräfte zusammengehalten, daß sie bei Zimmertemperatur in allen Flüssigkeiten *unlöslich* sind. Während man bei der Auflösung niedermolekularer Stoffe *Sättigungskonzentrationen* in allen Größenordnungen findet, gehen Hochpolymere entweder *ganz* oder *gar nicht* in Lösung. Der Vorgang beginnt mit einer *Quellung*, einer durch die Aufnahme von Lösungsmittel hervorgerufenen Volumenvergrößerung.

6.5.9.1.1. *Begrenzte und unbegrenzte Quellung.* Auch wenn genügend Lösungsmittel zur Verfügung steht, muß die Quellung nicht unbedingt zur Auflösung führen. Eine *begrenzte Quellung* zeigen vernetzte Polymere. In diesem Fall lagern sich um so mehr Lösungsmittelmoleküle in das Netzwerk ein, je weiter die Maschen sind, d. h., je geringer der Vernetzungsgrad ist. Doch auch bei Thermoplasten tritt manchmal nur begrenzte Quellung auf. Das bedeutet, daß die Wechselwirkungen zwischen Lösungsmittel und Makromolekülen die Anziehungskräfte der Makromoleküle untereinander nicht völlig überwinden können.
Bei *unbegrenzter Quellung* wird Auflösung erreicht. Es kann jedoch Tage dauern, bis eine homogene Lösung entstanden ist, auch wenn gerührt wird.

6.5.9.1.2. *Beeinflussung des Quellungsgrads.* Die bei einem bestimmten Polymer erzielbare Quellung hängt u. a. von der Beschaffenheit des Quellungsmittels ab. Eine Änderung des Quellungsgrads bedeutet Änderung des Volumens. Bei geeigneter Form oder bestimmtem Aufbau der Probe kann sich diese *Volumenänderung* hauptsächlich als *Längenänderung* äußern. Dabei kann z. B. ein einseitig befestigter Streifen ein angehängtes Gewicht heben, also Arbeit leisten. Solche künstlichen Muskeln, die chemische Energie direkt in mechanische Arbeit verwandeln, sind etwa seit 1950 bekannt. Sie sprechen entweder auf eine *Änderung des pH-Werts* oder des *Redoxpotentials* an.

6.5.9.2. Eigenschaften der Lösungen hochmolekularer Stoffe. Schon ein Zusatz von wenigen Prozent eines Polymers zu einem Lösungsmittel ruft eine außerordentlich große *Viskositäts-*

erhöhung hervor. Je besser das Lösungsmittel ist, um so mehr entfalten sich die Makromoleküle, um so größer ist die von ihnen ausgehende Behinderung der Fließvorgänge und um so höher liegt folglich die Viskosität der Lösungen. Sie hängt auch von der Größe der Makromoleküle ab, so daß Viskositätsmessungen einen Weg zur *Molmassenbestimmung* eröffnen (↑ 5.4.3.5.1.). In der Lackindustrie bezeichnet man dagegen bisweilen die Lösungsmittel als gut, mit denen Lösungen geringerer Viskosität, also letztlich höherer Konzentration, hergestellt werden können. Die so festgelegte Rangordnung ist der nach der thermodynamischen Güte der Lösungsmittel (↑ 5.4.3.4.5.) gerade entgegengerichtet.

6.5.9.2.1. *Filmbildung.* Beim Eintrocknen bleiben die gelösten makromolekularen Stoffe oft als glatte, durchsichtige Haut zurück. Bestimmte Folien und die Unterlage der Fotofilme werden technisch aus Lösungen gegossen.

Die beim Trocknen entstehenden Filme haften unter bestimmten Bedingungen fest an der Unterlage. Sie ergeben so einen *Lacküberzug* oder eine *Klebeverbindung* zwischen zwei festen Körpern.

6.5.9.2.2. *Lösungsmittelretention.* Die Trocknung eines aus Lösung erhaltenen Polymers kann äußerst langwierig werden: Das Lösungsmittel wird vom Polymer festgehalten und diffundiert nur langsam aus. Selbst in Folien sind nach tagelangem Trocknen im Vakuum noch Lösungsmittelspuren nachzuweisen.

6.6. Grenzflächen und disperse Systeme

6.6.1. Charakteristische Merkmale der Phasengrenzen

Alle Flüssigkeiten und festen Stoffe haben Oberflächen bzw. Grenzflächen. Als *Oberfläche* bezeichnet man die *Phasengrenze* zwischen einem kondensierten, d. h. flüssigen oder festen, Stoff und einer Gasphase. Den Begriff *Grenzfläche* bezieht man dagegen auf die Trennfläche zwischen kondensierten Phasen; im übergeordneten Sinne wird er auch verwendet, wenn allgemein die Begrenzungsflächen beliebiger Phasen (↑ 6.1.1.1.) gemeint sind. Phasengrenzen sind nicht nur die sichtbaren äußeren Grenzflächen von Flüssigkeiten und festen Stoffen. Es gibt sehr viele Stoffsysteme, die auch in ihrem Inneren mitunter sehr ausgedehnte Phasengrenzen aufweisen. Dazu gehören z. B. die Wände der oft submikroskopischen Porenkanäle bzw. Hohlräume in porösen und porigen Feststoffen, viele Korngrenzen in polykristallinen metallischen Werkstoffen oder Gesteinen, die „Oberflächen" der einzelnen Fäden oder Fasern in textilen Materialien und Papier oder die äußeren Begrenzungen der fein verteilten Tröpfchen und Partikel in Emulsionen bzw. Suspensionen. Gerade diese inneren Phasengrenzen bestimmen häufig wesentliche Eigenschaften der betreffenden Stoffsysteme.

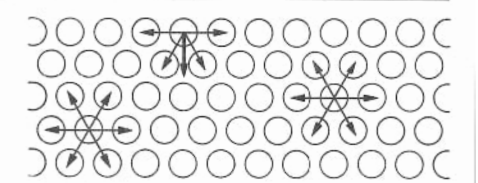


Abb. 6.6.1-1 Schematische Darstellung der zwischenmolekularen Wechselwirkungen im Inneren und an der Oberfläche einer kondensierten Phase

Phasengrenzen erstrecken sich immer über eine Tiefe von mehreren Atom- oder Moleküldurchmessern; die Ober- oder *Grenzflächen* sind daher in Wirklichkeit *Phasengrenzschichten.* Innerhalb dieser ändern sich die physikalischen und chemischen Eigenschaften der stofflichen Materie. Die Grenzschichten haben einen unterschiedlichen strukturellen Aufbau gegenüber dem Inneren der beiden aneinandergrenzenden Phasen und weisen oft auch noch eine unterschiedliche chemische Zusammensetzung auf. Die Teilchen in der Grenzschicht unterliegen *nicht* allseitig gleichmäßigen Kraftwirkungen wie im Phaseninneren (Abb. 6.6.1-1), was dort ein Minimum der potentiellen Wechselwirkungsenergie zur Folge hat. Phasengrenzen sind als Ganzes grundsätzlich energiereichere Gebiete in stofflichen Systemen.

Zahlreiche Vorgänge und Erscheinungen in Natur und Technik werden stärker von den Grenzflächen bestimmt als durch die Eigenschaften der beteiligten Stoffe an sich. Solche *Grenzflächenerscheinungen* und *-vorgänge* sind z. B. Gleiten, Reiben, Spalten, Schleifen, Polieren, Zerspanen, Mahlen, Sintern, Benetzen, Kleben, Löten und Schweißen, Drucken, Schreiben, Waschen, Färben, aber auch alle Formen von Niederschlagsbildung in der Atmosphäre, Vorgänge im Erdboden und viele Stoffwechselvorgänge in lebenden Organismen.

Die Ursache vieler Grenzflächenerscheinungen kennzeichnet man oftmals durch den Begriff „Wirken von Oberflächenkräften". Natürlich sind diese *Oberflächenkräfte* nichts anderes als die generell zwischen den Bausteinen der Stoffe wirkenden chemischen Bindungskräfte (↑ 5.1.1.) oder zwischenmolekularen Wechselwirkungskräfte (↑ 5.1.4.).

6.6.2. Zustandseigenschaften der Phasengrenzen

Wegen des Vorhandenseins von Phasengrenzen muß die gesamte Energie stofflicher Systeme immer auch einen Anteil an *Oberflächen-* oder

Grenzflächenenergie enthalten. Solange kompakte Phasen betrachtet werden, kann dieser Anteil wegen der verschwindend geringen Zahl von Grenzflächenteilchen gegenüber der Gesamtzahl der Teilchen im allgemeinen vernachlässigt werden. Wenn jedoch speziell der Einfluß oder die Eigenschaften der Phasengrenzschichten selbst betrachtet werden sollen oder wenn Stoffe mit stark vergrößerten Oberflächen vorliegen, z. B. als feinste Pulver oder als Nebeltröpfchen, ist es erforderlich, spezielle Zustandsgrößen (↑ 6.1.2.1.) zu berücksichtigen.

6.6.2.1. Thermodynamische Beschreibung. Von den Phasengrenzschichten sind im allgemeinen weder die Dicke noch die makroskopischen Eigenschaften exakt meßbar oder anderweitig bestimmbar, zumal sich letztere in ihnen stetig verändern. Zur thermodynamischen Erfassung der Phasengrenzen kann man das reale Grenzflächensystem durch ein Modellsystem ersetzen, in dem alle Besonderheiten und Abweichungen der Phasengrenze gegenüber den beiden Phaseninneren in eine fiktive dritte „Phase", die *Grenzflächenphase*, verlagert sein sollen. Dann läßt sich jede extensive Eigenschaft des Gesamtsystems additiv aus den betreffenden Eigenschaften der gewissermaßen als grenzflächenlos angesehenen Nachbarphasen α und β und der zweidimensionalen Grenzflächenphase (Kennzeichnung ihrer Eigenschaften durch ein hochgestelltes σ) darstellen, z. B. innere Energie $u = u^\alpha + u^\beta + u^\sigma$. Die so definierten grenzflächenenergetischen Zustandsgrößen heißen *gesamte Grenzflächenenergie* u^σ, *Grenzflächenenthalpie* h^σ, *Grenzflächenentropie* s^σ, *freie Grenzflächenenergie* f^σ und *freie Grenzflächenenthalpie* g^σ.

Zur Beschreibung beliebiger Zustandsänderungen unter Einbeziehung der Phasengrenzen können auf das Ersatzsystem alle bekannten thermodynamischen Beziehungen und Zusammenhänge angewandt werden. Die durch eine differentielle Zustandsänderung hervorgerufene Änderung z. B. der freien Enthalpie läßt sich dann beschreiben durch $dg = dg^\alpha + dg^\beta + dg^\sigma$. Für die Grenzflächenphase entfällt, da sie laut Definition kein Volumen einnimmt, die partielle Abhängigkeit vom Druck p; dafür muß die Abhängigkeit von der zweidimensionalen Ausdehnung A berücksichtigt werden. Damit ergibt sich für das Gleichgewicht $\mu_i^\alpha = \mu_i^\beta = \mu_i^\sigma$ (↑ 6.1.2.4.) die *Gibbssche Fundamentalgleichung für die Grenzflächenphase*

$$(1) \quad dg^\sigma = -s^\sigma\, dT + (\partial g^\sigma/\partial A)_{T,p,n_i}\, dA + \sum \mu_i\, dn_i^\sigma.$$

μ_i ist das chemische Potential, n_i^σ die Stoffmenge der Komponente i in der Grenzflächenphase. Mit $v^\alpha + v^\beta = v$, $s^\alpha + s^\beta + s^\sigma = s$ und $dn_i^\alpha + dn_i^\beta + dn_i^\sigma = dn_i$ gilt dann für das Gesamtsystem $dg = v\, dp - s\, dT + (\partial g^\sigma/\partial A)_{T,p,n_i}\, dA + \sum \mu_i\, dn_i$. Der partielle Differentialquotient $(\partial g^\sigma/\partial A)_{T,p,n_i}$ wird als *spezifische freie Grenzflächenenthalpie* G^σ bezeichnet. Er gibt die Änderung der freien Enthalpie des Gesamtsystems bei Änderung der Grenzflächenausdehnung um die Flächeneinheit an, wenn dabei der Druck, die Temperatur und die stoffliche Zusammensetzung unverändert bleiben.

Analog zu G^σ lassen sich auch die *spezifische Grenzflächenenthalpie* H^σ, die *spezifische Grenzflächenentropie* S^σ usw. definieren. Alle aus der Kombination der beiden Hauptsätze der Thermodynamik ableitbaren Beziehungen gelten auch für die Grenzflächenphase und beschreiben den Grenzflächeneinfluß auf stoffliche Zustände und Zustandsänderungen.

6.6.2.2. Oberflächeneigenschaften reiner Stoffe

6.6.2.2.1. *Spezifische freie Grenzflächenenthalpie und Oberflächenspannung.* Die *Oberflächenspannung* σ ist definiert als die zur isotherm-isobaren und reversiblen Erzeugung der Flächeneinheit neuer Oberflächen notwendige mechanische Arbeit: $(dw_{rev}^A)_{T,p} = (dg^\sigma)_{T,p} = (d(G^\sigma A))_{T,p} = \sigma\, dA$. Die Stoffkonstante σ hängt von der Temperatur T und dem Druck p ab. Ihre Dimension ist Energie/Fläche bzw. Kraft/Länge; die heute oft verwendete Einheit ist mNm^{-1} ($= mJm^{-2} = gs^{-2}$), wodurch die Zahlenwerte die gleichen wie bei den früher verwendeten Einheiten $erg \cdot cm^{-2} = dyn \cdot cm^{-1}$ bleiben.

Oberflächenvergrößerung bedeutet bei Flüssigkeiten, daß Teilchen aus dem Phaseninneren in die Oberfläche, also in einen höheren Energiezustand, gebracht werden. Die Oberflächenspannung ist daher immer positiv. Infolge der hohen Beweglichkeit der Bausteine ordnet sich jedes neu in die Oberfläche gelangende Teilchen innerhalb von Nanosekunden strukturell so an, daß sein Energiezustand dem Gleichgewichtszustand entspricht. Bei *reinen Flüssigkeiten* ist G^σ unabhängig von A und damit $\sigma = G^\sigma$.

Aus der Sicht der Bausteine umfaßt eine Oberflächenvergrößerung stets zwei Teilvorgänge, die nur bei Flüssigkeiten durch die hohe Teilchenbeweglichkeit „automatisch" miteinander gekoppelt sind: Überführung in die Grenzfläche und strukturelle Umordnung zum energieärmsten Zustand. Die Bausteine in festen Stoffen sind nur wenig beweglich; deswegen kann hier im allgemeinen nur der erste Teilvorgang vonstatten gehen (z. B. beim Zerbrechen eines Festkörpers). Andererseits ist bei vielen festen Stoffen eine Oberflächenvergrößerung möglich, ohne daß überhaupt Bausteine aus dem Inneren in die Grenzfläche überführt werden (z. B. Dehnung, Streckung elastischer Stoffe). Hierbei erfolgt nur eine Veränderung der gegenseitigen Teilchenabstände und gegebenenfalls der strukturellen Anordnung. Bei *reinen festen Stoffen* ist daher die spezifische

freie Genzflächenenthalpie nicht unabhängig von der Oberflächengröße A. Es gilt vielmehr $\sigma = G^\sigma + A(\partial G^\sigma/\partial A)_{T,p}$.

6.6.2.2.2. *Oberflächenspannungen reiner Flüssigkeiten.* Oberflächenspannungen von Flüssigkeiten können direkt experimentell bestimmt werden. Im normalen Druck- und Temperaturbereich liegen sie je nach den zwischenmolekularen Wechselwirkungskräften meist in der Größenordnung von etwa 20 bis 40 mNm^{-1} (Tab. 6.6.2-1).

Tab. 6.6.2-1 Oberflächenspannungen σ (in mNm^{-1}) und Temperaturkoeffizienten $d\sigma/dT$ (in $\text{mNm}^{-1}\text{K}^{-1}$) verschiedener Flüssigkeiten bei 20 °C

Flüssigkeit	σ	$-d\sigma/dT$
H_2O	72,82	0,1571
CS_2	32,4	0,152
C_2H_5OH	22,3	0,086
$n\text{-}C_4H_9OH$	24,7	0,083
$C_2H_5OC_2H_5$	17,0	0,116
HCOOH	37,6	0,114
CH_3COOH	27,7	0,097
CCl_4	26,7	0,119
C_6H_6	29,2	0,127
$C_6H_5CH_3$	28,4	0,111
$C_6H_5NO_2$	43,4	0,113
$i\text{-}C_8H_{18}$	19,1	0,087

Die Ausnahmestellung des Wassers erklärt sich aus den starken Wasserstoffbrückenbindungen (↑ 6.3.2.1.2.), auch wenn man berücksichtigt, daß wegen der Kleinheit der Moleküle bei einer Oberflächenvergrößerung um 1 cm² bedeutend mehr Teilchen in den Grenzflächenzustand überführt werden müssen als bei den meisten anderen Flüssigkeiten.

Geschmolzene Salze haben bei Temperaturen zwischen 400 und 1000 °C Oberflächenspannungen von 50 bis 300 mNm^{-1}, geschmolzene Silicatgläser (1000 bis 1500 °C) von 200 bis 400 mNm^{-1} und geschmolzene Metalle (400 bis 1600 °C) von 400 bis 2000 mNm^{-1}. Für Quecksilber ist bei 25 °C $\sigma = (475 \pm 10)\,\text{mNm}^{-1}$ und $d\sigma/dT = -0{,}224\,\text{mNm}^{-1}\text{K}^{-1}$.

Temperaturerhöhung bedeutet immer Vergrößerung der Teilchenenergie und damit Verringerung des Energiebedarfs, um Teilchen in den Oberflächenzustand zu überführen; die Oberflächenspannung nimmt grundsätzlich mit steigender Temperatur ab.

6.6.2.2.3. *Struktur der Flüssigkeitsoberflächen.* Durch Untersuchung von Lichtreflexionen ist auch experimentell nachgewiesen, daß die Oberflächen ruhender Flüssigkeiten nahezu ideal glatt sind. Die Phasengrenzschicht erstreckt sich nur über einen Bereich von ein bis zwei Teilchendurchmessern. Mit steigender Temperatur nimmt die „Schärfe“ der Phasengrenze ab.

Auch bei normaler Temperatur herrscht jedoch in der Grenzschicht eine starke Teilchenfluktuation. Im Verdampfungsgleichgewicht (↑ 6.3.2.3.1.) gehen ständig Teilchen aus der flüssigen Phase in den Dampfraum und umgekehrt über, z. B. bei Wasser von 20 °C etwa $3 \cdot 10^{20}$ Moleküle/cm²s oder etwa $3 \cdot 10^6$ Moleküle/s durch die Fläche, die ein einziges Wassermolekül in der Oberfläche einnimmt. Wegen der intensiven Bewegung und der geringen „Sprungweite“ der Moleküle (↑ 6.3.5.3.) erfolgt außerdem ein noch stärkerer ständiger Austausch von Teilchen zwischen der äußersten Grenzschicht und dem Flüssigkeitsinneren. Die mittlere Verweilzeit eines bestimmten Teilchens in der Oberfläche liegt daher nur in der Größenordnung einer Mikrosekunde. Trotz dieser ungeheuren „Unruhe“ in einer Flüssigkeitsoberfläche sind aber die zwischenmolekularen Wechselwirkungskräfte immer stark genug, um die Oberfläche nicht zu einem verschwommenen Nebelschleier werden zu lassen.

6.6.2.2.4. *Struktur der Oberflächen fester Stoffe.* Bei festen Stoffen gibt es niemals ideal glatte Oberflächen, auch nicht bei Kristallen, die unter weitgehend idealen Wachstumsbedingungen gezüchtet wurden. Neben z. B. Stufen und Kanten sind stets Fehlstellen in der Gitteranordnung der Bausteine vorhanden, vor allem als Austrittsstellen von Versetzungen oder von Korngrenzen (↑ 6.4.7.3.4.). Weitere Störungen werden durch die stets vorhandenen Fremdbausteine (Verunreinigungen) verursacht. Alle diese *Unregelmäßigkeiten* sind nicht nur *struktureller* und *stofflicher Natur*, sondern wirken sich grundsätzlich auch in einer *energetischen Heterogenität* aus. Reale Festkörper, z. B. technische Werkstoffe, haben meist stark unregelmäßig gestaltete, zerklüftete und rauhe Oberflächen. Selbst mit den aufwendigsten Methoden der Oberflächenfeinbearbeitung ist es z. B. bei Metallen kaum möglich, unter Rauhtiefen von 10^{-1} bis 10^{-2} µm zu kommen, was im Vergleich zu den atomaren Dimensionen von 10^{-4} bis 10^{-3} µm noch sehr grob ist. Daher sind bei festen Stoffen die wahren äußeren Oberflächen auch meist beträchtlich größer als die geometrischen. Die atomaren Bausteine in Ecken, Kanten, Rauheitsspitzen und -tälern befinden sich in stark unterschiedlichen Energiezuständen. Diese wirken auch auf die Nachbarbausteine in Richtung Phaseninneres, so daß sich die Oberflächenunregelmäßigkeiten stets auch mehr oder weniger tief ins Innere fortsetzen. Damit sind die Phasengrenzschichten bei festen Stoffen im allgemeinen nicht nur wesentlich dicker als bei Flüssigkeiten, sondern

auch von Oberflächenstelle zu Oberflächenstelle verschieden.

Eine besonders tiefe Ausdehnung hat die Grenzschicht immer dann, wenn der Festkörper mechanisch bearbeitet wurde. Durch Drehen, Fräsen, Schleifen oder Polieren z. B., aber auch durch Walzen, Pressen oder Ziehen, werden zwar die gröbsten Stufen und Spitzen eingeebnet, es entstehen dabei aber immer Oberflächenschichten, in denen die kristallographische Ordnung und das Kristallgefüge stark gestört, „verschmiert" sind und die häufig durch Einlagerung von Bestandteilen aus den angrenzenden Phasen, z. B. aus den Poliermitteln oder der Luft, verunreinigt sind. Bei metallischen Werkstoffen führt diese submikrokristalline *Beilbyschicht* (Abb. 6.6.2-2) häufig zu einer Oberflächen-(„Kalt"-)Verfestigung. Besondere Verhältnisse liegen auch vor, wenn die festen Oberflächen Mikrorisse oder feine Poren aufweisen (s. a. 6.4.7.3.)

6.6.2.2.5. *Oberflächenenergien fester Stoffe.* Eine sowohl für die Bestimmung als auch für die Interpretation von Oberflächenenergien von Festkörpern bestehende Schwierigkeit liegt darin, daß sich diese makroskopisch meist aus zwei Anteilen zusammensetzen: Einen durch den eigentlichen „Grenzflächenzustand" der betreffenden Bausteine verursachten Anteil und einen zweiten, der auf mechanische Spannungen in der Grenzschicht zurückgeht. Hierin liegt auch eine der Hauptursachen dafür, daß die auf verschiedenen experimentellen Wegen ermittelten *Oberflächenspannungen von Festkörpern* (Tab. 6.6.2-3) im allgemeinen nur als Näherungswerte angesehen werden können. Man bezeichnet zusammenfassend alle Stoffe mit Oberflächenspannungen über etwa 100 mNm^{-1} als *Festkörper bzw. Feststoffe mit hochenergetischen Oberflächen.* Alle Stoffe, in denen der Zusammenhalt der Bausteine im wesentlichen durch

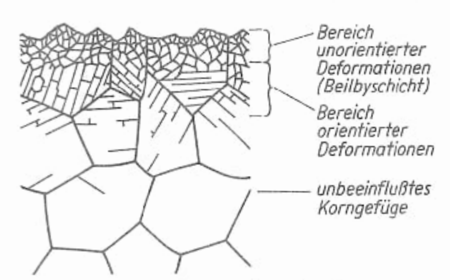


Abb. 6.6.2-2 Gefügestruktur im Oberflächenbereich eines metallischen Werkstoffs nach einer mechanischen Bearbeitung (schematisch)

die schwächeren zwischenmolekularen Wechselwirkungskräfte erfolgt und die demzufolge Oberflächenspannungen unter etwa 100 mNm^{-1} besitzen, sind dagegen *Festkörper bzw. Feststoffe mit niederenergetischen Oberflächen.* Die σ-Werte solcher festen Stoffe, wie der organischen hochpolymeren Natur- und Kunststoffe, sind auf einem indirekten Wege zugänglich (s. a. 6.6.6.3.).

6.6.2.3. Gekrümmte Grenzflächen. Ohne äußere Einwirkungen, z. B. durch Gravitationskräfte oder durch Gefäßwände, nehmen Flüssigkeiten stets eine *Kugelform* an (Regentröpfchen, Seifenblasen, Gasblasen in Flüssigkeiten). Darin drückt sich das Streben nach einem möglichst energiearmen Zustand aus: Von allen geometrischen Körpern hat die Kugel die kleinste Oberfläche im Verhältnis zum Volumen. Die weitaus häufiger anzutreffenden *ebenen Flüssigkeitsoberflächen* stellen dagegen immer eine Zwangsform unter der Wirkung der Schwerkraft dar, die das Bestreben zur Minimalisierung der Oberfläche bei weitem überkompensiert.

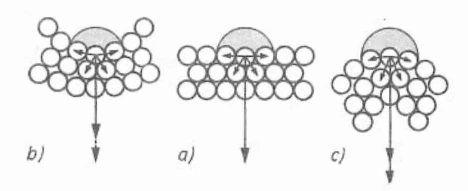


Abb. 6.6.2-4 Schematische Darstellung der Grenzflächenkräfte; a) bei ebenen, b), c) bei gekrümmten Flüssigkeitsoberflächen

Bei *gekrümmten Oberflächen* sind die Teilchen in der Phasengrenze mehr (Abb. 6.6.2-4b) oder weniger (Abb. 6.6.2-4c) von nächsten Nachbarn umhüllt, können also im Vergleich zur ebenen Oberfläche (Abb. 6.6.2-4a) einen größeren oder kleineren Teil ihrer „freien Oberflächenkräfte" (schraffierte Flächen) absättigen. Dem „normalen" Zug auf diese Teilchen in Richtung

Tab. 6.6.2-3 Oberflächenspannungen σ (in mNm^{-1}) verschiedener polykristalliner Festkörper bei den Temperaturen ϑ (in °C)

Festkörper	ϑ	σ
Al	150 bis 210	1140 ± 200
Cu	950 bis 1050	1710 ± 100
Ag	875 bis 930	1140 ± 90
Au	1020 bis 1045	1400 ± 65
Ni	1250 bis 1450	1820 ± 180
Fe	1360 bis 1400	2150 ± 320
Cr	1400 bis 1700	2200 ± 250
Pt	1310	2340 ± 800
W	1750	2900 ± 300
NaCl	25	345
CaO	25	1540
Glimmer	25	310

Phaseninneres (↑ Abb. 6.6.1-1) überlagert sich eine negative oder positive Zusatzkraft. Die Moleküle im Inneren unterliegen gegenüber der ebenen Oberfläche einem kleineren oder größeren Druck. Diese „Zusatzdrücke" äußern sich als ***Druckdifferenz*** beiderseits der Phasengrenze.

6.6.2.3.1. *Thermodynamische Beschreibung.* Bei gekrümmten Grenzflächen ist der Druck auf der konkaven Seite p_{kk} immer größer als auf der konvexen p_{kv}. Die Druckdifferenz $p_{kk} - p_{kv}$ hängt von der Oberflächenspannung σ ab, denn diese ist ihre Ursache. Sie ist außerdem um so größer, je kleiner der Krümmungsradius, je stärker die Krümmung ist.

Ausgehend von der Gibbsschen Fundamentalgleichung für die Grenzflächenphase (↑ 6.6.2.1.) ergibt sich für beliebig gekrümmte Grenzflächen mit den beiden Hauptkrümmungsradien r_1 und r_2 die wichtige *Young-Laplacesche Gleichung* $(p_{kk} - p_{kv}) = \Delta p_k = \sigma\left(\frac{1}{r_1} + \frac{1}{r_2}\right)$. Für den Spezialfall einer kugelförmigen Krümmungsfläche ($r_1 = r_2 = r$) folgt für den *Krümmungs-* oder *Kapillardruck* Δp_k die in der Praxis meist anwendbare Form $\Delta p_k = 2\sigma/r$.

Zusätzlich zu irgend einem äußeren Druck, z. B. dem hydrostatischen Druck einer Flüssigkeit über einer Gasblase, herrscht also im Inneren einer solchen Blase ein Überdruck Δp_k. Auch das Innere eines Flüssigkeitströpfchens vom Radius r steht unter diesem Überdruck. Für $r \to \infty$, d. h. über einer ebenen Oberfläche, wird Δp_k dagegen Null; im thermodynamischen und mechanischen Gleichgewicht herrscht beiderseits einer ebenen Phasengrenze *kein* Druckunterschied. Der Kapillardruck wird allerdings immer erst bei sehr starker Krümmung bedeutungsvoll. So ist z. B. in einem Wassertröpfchen mit einem Durchmesser von $2r = 1$ mm $\Delta p_k \approx 3$ mbar; bei $2r = 1$ µm wird $\Delta p_k \approx 3$ bar und bei $2r = 10$ nm $\Delta p_k \approx 300$ bar.

6.6.2.3.2. *Dampfdrücke über gekrümmten Oberflächen.* Ein erhöhter Druck im Inneren von Tröpfchen bewirkt auch einen höheren Dampfdruck (↑ 6.3.2.3.). Er wird um so größer, je größer die Oberflächenspannung σ der Flüssigkeit und je kleiner der Krümmungsradius r ist. Aus der Gibbsschen Fundamentalgleichung für die Grenzflächenphase läßt sich für kugelförmige Flüssigkeitstropfen ableiten: $RT \ln (p/p_0) = (2\sigma V)/r$. In dieser *Kelvingleichung* bedeuten

Tab. 6.6.2-5 Anwendung der *Kelvingleichung* auf Wassertröpfchen bei $T = 293$ K; $\sigma = 72{,}8$ mNm^{-1}, $V = 18$ cm^{3}mol^{-2}, $p_0 = 23{,}3 \cdot 10^1$ Pa

r in µm	10^0	10^{-1}	10^{-2}	10^{-3}
p/p_0	1,001	1,011	1,114	2,95

p Dampfdruck der Tröpfchen vom Radius r, p_0 Dampfdruck über der ebenen Flüssigkeitsoberfläche bei der gleichen Temperatur T, R molare Gaskonstante und V molares Volumen der Flüssigkeit. Mit formal negativem Vorzeichen von r gilt die Kelvingleichung für Dampfbläschen in Flüssigkeiten. Wie Tab. 6.6.2-5 am Beispiel von Wassertropfen zeigt, wird die Dampfdruckzunahme erst bei sehr kleinen Krümmungen merklich.

Die Dampfdruckzunahme oder -abnahme ist von großer praktischer Bedeutung, z. B. für die spontane Keimbildung bei der Kondensation von übersättigtem Dampf, für die Vereinigung kleinster Flüssigkeitströpfchen zu größeren durch isotherme Destillation (Abregnen von Nebel), für die Keimbildung beim Sieden überhitzter Flüssigkeiten oder für die Erscheinung der *Kapillarkondensation.* Beim Eindringen von Dämpfen in enge, kapillare Hohlräume, z. B. hochporöser Materialien, in denen die betreffende Flüssigkeit konkave Menisken ausbildet, kann der Dampfdruck soweit absinken, daß eine Kondensation auch oberhalb der normalen Kondensationstemperatur erfolgt. In der Technik nutzt man diesen Vorgang z. B. zur Wiedergewinnung von Lösungsmitteldämpfen.

Auch sehr kleine Kristalle haben einen höheren Dampfdruck und damit z. B. einen niedrigeren Schmelzpunkt als der kompakte Stoff. Wegen ihrer größeren freien Oberflächenenthalpie haben feste Stoffe in feinster Verteilung andererseits größere Verbrennungs- oder Lösungswärmen sowie erhöhte Löslichkeiten. Das Bestreben zur Verringerung der freien Oberflächenenthalpie führt bei sehr fein verteilten festen Stoffen zu einer Teilchenvergrößerung. Dieser Vorgang ist z. B. bei heterogenen Katalysen, für Rekristallisations- und Sintervorgänge von Bedeutung.

6.6.2.3.3. *Erscheinungen in Kapillaren.* Der Krümmungsdruck Δp_k ist auch die ***Triebkraft*** für das Eindringen von Flüssigkeiten in poröse Stoffe, wenn in deren Hohlräumen konkave Menisken ausgebildet werden. Beispiele hierfür sind das Aufsaugen von Flüssigkeiten durch Schwämme oder ähnlich aufgebaute Materialien, das „Wegschlagen" von Tinte oder Druckfarbe auf Papier, das Hochsteigen von Wasser in porösem Mauerwerk oder das Aufsteigen von Öl oder flüssigem Paraffin in Lampen- bzw. Kerzendochten.

In einer engen Kapillare steigt eine Flüssigkeit bei Ausbildung eines konkaven Meniskus soweit über das Niveau der äußeren ebenen Flüssigkeitsoberfläche hoch, bis dem Krümmungsdruck durch das Gewicht der Flüssigkeitssäule die Waage gehalten wird (Abb. 6.6.2-6a). Für die *Steighöhe h* in einer Kapillare vom Radius r_k gilt im Gleichgewicht $2\sigma/r_k$

$= \varrho gh$, σ ist die Oberflächenspannung, ϱ die Dichte der Flüssigkeit; g die Erdbeschleunigung. Die Größe $2\sigma/g\varrho = r_k h = a^2$ wird als *Kapillarkonstante* bezeichnet. Die Ausbildung eines konkaven Meniskus ist gleichbedeutend mit der Eigenschaft der Flüssigkeit, die Kapillarwand

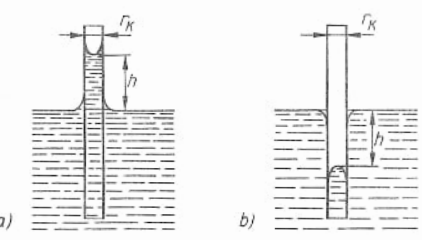


Abb. 6.6.2-6 a) Aszension benetzender, b) Depression nicht benetzender Flüssigkeiten in Kapillaren

zu benetzen (↑ 6.6.6.). Es gibt aber auch Systeme, in denen die Flüssigkeit die Kapillarwand nicht benetzt; dann bildet sie in engen Kapillaren einen konvexen Meniskus aus. Wegen des dann entgegengesetzten Vorzeichens von Δp_k ergibt sich eine *Kapillardepression*, z. B. bei Hg in Glaskapillaren (Abb. 6.6.2-6b).
Von praktischer Bedeutung ist auch die *Steiggeschwindigkeit* benetzender Flüssigkeiten in Kapillaren. Laminare Strömung vorausgesetzt, gilt hierfür $dh_t/dt = r_k^2/(8\eta h_t)\,[(2\sigma/r_k) - \varrho g h_t]$. Die Einstellgeschwindigkeit hängt nicht nur von der Viskosität η der Flüssigkeit, sondern auch von der jeweiligen, zur Zeit t bereits erreichten Höhe h_t ab. Für die Bewegung in einer horizontalen Kapillare vereinfacht sich diese Beziehung, da die durch den Kapillardruck gegebene treibende Kraft dann konstant bleibt, zu der *Washburngleichung* $h_t = \sqrt{(\sigma r_k t)/(2\eta)}$.

6.6.3. Allgemeines zu Wechselwirkungen an Phasengrenzen

Einem möglichst energiearmen Zustand können sich Grenzflächen außer durch Verkleinerung auch dadurch annähern, daß sie mit Teilchen der benachbarten Phasen in Wechselwirkung treten. Die dabei betätigten Bindungs- oder anderen Wechselwirkungskräfte reduzieren die in der Grenzschicht enthaltene Energie. Dies ist die *Grundursache* für alle *Vorgänge*, die ablaufen, wenn Bausteine zweier Phasen miteinander in Kontakt kommen, wie auch für alle *Erscheinungen*, die an den Grenzflächen dabei auftreten.

6.6.3.1. Grenzflächenspannungen

6.6.3.1.1. *Phänomenologische Aussagen.* Unter der *Grenzflächenspannung* $\gamma^{\alpha\beta}$ versteht man die Arbeit, die zur isotherm-isobaren und reversiblen Vergrößerung der Grenzfläche zwischen zwei kondensierten Phasen α und β um die Flächeneinheit aufgewandt werden muß.
Grenzflächenspannungen zwischen zwei Flüssigkeiten sind experimentell bestimmbar. Wie die Beispiele in Tab. 6.6.3-1 zeigen, sind sie grundsätzlich kleiner als die größere der Oberflächenspannungen beider Flüssigkeiten, oft sogar kleiner als die niedrigere von beiden. Zwischen Wasser und organischen Flüssigkeiten sind die γ-Werte um so kleiner, je größer die gegenseitige Löslichkeit ist.
Grenzflächenspannungen zwischen Flüssigkeiten und festen Stoffen sind grundsätzlich kleiner als die Oberflächenspannung des Festkörpers bzw. bei Feststoffen mit niederenergetischen Oberflächen als die größere der beiden Oberflächenspannungen. Für Systeme mit hochenergetischen festen Stoffen gibt es praktisch noch keine quantitativen Angaben, wogegen bei niederenergetischen festen Oberflächen theoretische und experimentelle Abschätzungen möglich sind (↑ 6.6.6.3.).

Tab. 6.6.3-1 Grenzflächenspannungen γ (in mNm^{-1}) zwischen Wasser ($\sigma_W = 72{,}8\ mNm^{-1}$) und verschiedenen organischen Flüssigkeiten (σ_{OF} in mNm^{-1}) sowie Löslichkeiten c_s (in Masse-%) dieser Flüssigkeiten in Wasser bei 20 °C

System	σ_{OF}	$\gamma_{W/OF}$	c_s
W/höhere Alkane	24 bis 28	≈ 54	
W/n-C_8H_{18}	21,8	50,8	$7 \cdot 10^{-5}$
W/CCl_4	26,7	45,1	$8 \cdot 10^{-2}$
W/C_6H_6	29,2	35,0	$6 \cdot 10^{-2}$
W/$C_6H_5NO_2$	43,4	26,0	$2 \cdot 10^{-1}$
W/$C_2H_5OC_2H_5$	17,0	10,7	6,6
W/$CH_3COOC_2H_5$	23,9	2,9	7,9
W/n-C_4H_9OH	24,7	1,6	20,8

6.6.3.1.2. *Molekulartheoretische Betrachtungen.* Was den Zusammenhalt der Bausteine im Inneren von Flüssigkeiten und festen Stoffen bewirkt (↑ 5.1.3. und 5.1.4.), muß auch in den Oberflächenenergien zum Ausdruck kommen. Allgemein gilt daher, daß sich die Oberflächenspannungen *additiv* aus verschiedenen Wechselwirkungsenergieanteilen zusammensetzen; z. B. gilt bei Flüssigkeiten $\sigma = \sigma^d + \sigma^i + \sigma^\mu + \sigma^{\mu(h)}$. σ^d heißt *Dispersionsanteil*, σ^i *Induktionsanteil*, σ^μ *Orientierungsanteil* und $\sigma^{\mu(h)}$ *Wasserstoffbrückenanteil der Oberflächenspannung*.
Zwischen unterschiedlichen Stoffen können beim Zusammenbringen Wechselwirkungen über die Phasengrenze hinweg in der Regel nur über gleiche Arten von Bindungs- oder Wechselwirkungskräften erfolgen. Bei Kontakt zweier Phasen können daher immer nur die auf

die gleichen Kräftearten zurückgehenden Anteile an den Oberflächenenergien beider Partner kompensiert werden. Die Restanteile bleiben in der gebildeten Grenzfläche enthalten und ergeben die Grenzflächenspannung. Grundsätzlich immer können beliebige Phasen über die Dispersionskräfte in Wechselwirkung treten, während die Möglichkeit einer Betätigung und Absättigung der anderen Kräftearten nur bei chemischer Ähnlichkeit oder Entstehung einer Induktionswechselwirkung gegeben ist.

Bei Systemen aus einer völlig unpolaren Flüssigkeit α, bei der also $\sigma^\alpha = \sigma^{\alpha d}$ ist, und einer beliebigen anderen Flüssigkeit β ergibt sich daraus z. B. allgemeingültig $\gamma^{\alpha\beta} = \sigma^\alpha + \sigma^\beta - 2\Phi(\sigma^\alpha\sigma^{\beta d})^{1/2}$. $\sigma^{\beta d}$ ist der Dispersionsanteil der Oberflächenspannung des Stoffs β. Dispersionsanteile können so für alle Flüssigkeiten und sogar festen Stoffe aus Grenzflächenspannungsmessungen und Benetzungsuntersuchungen an geeigneten Stoffsystemen unter Verwendung weiterer, aus der Theorie folgender Zusammenhänge mit guter Näherung ermittelt werden. Φ ist ein Faktor, der aus den Wechselwirkungspotentialen für die gleichartigen und die unterschiedlichen Teilchenarten zugänglich ist. Damit sind auch die entsprechenden Grenzflächenspannungen zu berechnen. Die theoretischen Ansätze liefern darüber hinaus auch die Möglichkeit einer Abschätzung der polaren Anteile der Oberflächenspannungen von Flüssigkeiten und niederenergetischen festen Stoffen. So setzt sich z. B. die Oberflächenspannung von Wasser zusammen aus $\sigma^d_W = 19{,}7\ \mathrm{mNm^{-1}}$, $\sigma^i_W = 2{,}0\ \mathrm{mNm^{-1}}$ und $\sigma^{\mu(h)}_W = 51{,}1\ \mathrm{mNm^{-1}}$.

6.6.3.1.3. *Kohäsions- und Adhäsionsarbeit.* Unter *Adhäsion* versteht man die Wechselwirkung eines festen Stoffs mit größeren Partikeln oder kompakten flüssigen oder festen Phasen, die zu einer *Haftung* der Partikel oder Phasen an der festen Oberfläche führt. Der Begriff *Kohäsion* bezeichnet dagegen den Zusammenhalt der Teilchen im Inneren eines Stoffs. *Adhäsions-* und *Kohäsionskräfte* sind also in einer älteren Bezeichnungsweise die chemischen Bindungs- und zwischenmolekularen Wechselwirkungskräfte.

Die *Kohäsionsarbeit* ist definiert als die isotherm-isobare reversible Arbeit, die aufgewandt werden muß, um eine Flüssigkeitssäule von 1 cm² Querschnitt senkrecht zur Säulenachse in zwei Teile zu trennen. Hierbei werden 2 cm² neuer Oberfläche der Flüssigkeit α gebildet; die Kohäsionsarbeit w_{Ko} ist gleich der doppelten Oberflächenspannung σ^α der Flüssigkeit. Als *Adhäsionsarbeit* wird die isotherm-isobare reversible Arbeit definiert, die zur Trennung zweier verschiedener aneinandergrenzender Flüssigkeitssäulen von 1 cm² Querschnitt an der Grenzfläche erforderlich ist. Da hierbei je 1 cm² Oberfläche der beiden Flüssigkeiten α und β neu gebildet werden und 1 cm² Grenzfläche verschwindet, ist die Adhäsionsarbeit w_{Ad} gleich der Summe der Oberflächenspannungen σ^α und σ^β, vermindert um die Grenzflächenspannung $\gamma^{\alpha\beta}$ zwischen den beiden Flüssigkeiten.

Allgemein ergibt ein Vergleich der Kohäsionsarbeit einer Flüssigkeit mit der Adhäsionsarbeit gegenüber einer zweiten Flüssigkeit eine Vorstellung von der relativen Stärke der Anziehungskräfte zwischen den Teilchen der ersten Flüssigkeit gegenüber denen zwischen den unterschiedlichen Teilchen.

6.6.3.2. Orientierung in Phasengrenzen. Beim Kontakt zweier Stoffe können neben dem energetischen Effekt im molekularen Bereich strukturelle Veränderungen auftreten, wenn die Teilchen wenigstens einer Phase genügend beweglich sind. Dabei sind häufig die Moleküle nicht nur als Ganzes zu sehen, weil einzelne *Molekülteile* in unterschiedliche Wechselwirkungen mit Teilchen der jeweiligen Nachbarphase treten können. Bei längerkettigen Alkoholen $H_3C—(CH_2)_n—OH$ z. B. kann die Kohlenwasserstoffkette auf Nachbarteilchen nur Dispersionskräfte ausüben. Die Hydroxylgruppe dagegen ist bestrebt, mit ihrer Umgebung über die stärkeren Wasserstoffbrückenbindungen in Wechselwirkung zu treten, wodurch ein größerer Betrag an Oberflächenenergie abgebaut und ein energieärmerer Zustand erreicht wird. Analoges trifft auf alle anderen Moleküle mit unpolaren und polaren Molekülteilen, z. B. den Gruppen $—COOH$, $—OSO_3H$, $—SO_3H$, $—NH_2$, $=NH$ oder $—CONH—$, zu.

Ein solcher Molekülaufbau drückt sich bereits in bestimmten *Teilchenorientierungen* in der Oberfläche und damit in der Oberflächenspannung z. B. der erwähnten Alkohole aus (Abb. 6.6.3-2a), indem vorzugsweise die unpolaren Gruppen gegen die Gasphase gerichtet sind. An der Phasengrenze z. B. gegen Wasser

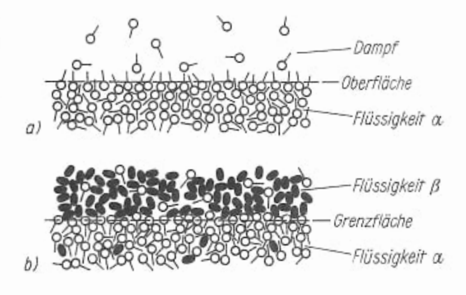


Abb. 6.6.3-2 Anordnung von Molekülen mit unpolaren und polaren Molekülteilen; a) in Flüssigkeitsoberflächen, b) in Flüssigkeitsgrenzflächen (schematisch)

sind die Alkoholmoleküle dagegen bevorzugt mit ihren Hydroxylgruppenenden den Wassermolekülen zugewandt (Abb. 6.6.3-2b).

6.6.4. Adsorption

Sie gehört zu den theoretisch wie praktisch bedeutsamsten Grenzflächenerscheinungen. Man versteht darunter die *Anreicherung von Teilchen* aus einer fluiden (gasförmigen oder flüssigen) Phase in einer Phasengrenze. Der Stoff, der in seiner Grenzfläche bestimmte Atome, Moleküle oder Ionen anreichert, heißt *Adsorbens* oder *Adsorptionsmittel*, z. B. Aktivkohle in einem Gasmaskenfilter. Die zur Adsorption fähige Teilchenart, z. B. H_2S in einer Luftatmosphäre, wird vor der Adsorption als *Adsorptiv*, im adsorbierten Zustand als *Adsorbat* bezeichnet. Das System Adsorbens + Adsorbat, im Beispiel die Aktivkohle mit adsorbiertem H_2S, nennt man *Adsorptionskomplex*. In vielen Fällen kann ein Adsorptionskomplex auch wieder in seine ursprünglichen Bestandteile zerlegt werden; der Vorgang Adsorbat → Adsorptiv wird als *Desorption* bezeichnet.

Nach der Stärke der Wechselwirkungskräfte im Adsorptionskomplex unterscheidet man zwischen physikalischer Adsorption und Chemisorption, die aber fließend ineinander übergehen. Sind allein zwischenmolekulare Wechselwirkungskräfte wirksam, dann liegt eine *physikalische Adsorption* (*Physisorption*) vor. Erreicht dagegen die Wechselwirkung Adsorbens-Adsorbat die Größenordnung chemischer Bindungskräfte, spricht man von einer *Chemisorption*. Sehr oft geht eine Chemisorption in eine echte Verbindungsbildung über, z. B. Oxidschichten auf Metalloberflächen.

Die physikalische Adsorption verläuft im allgemeinen ungehemmt und führt zur raschen Einstellung eines *Adsorptions-Desorptionsgleichgewichts*, was bei der Chemisorption häufig nicht der Fall ist.

6.6.4.1. Adsorption an Grenzflächen von Flüssigkeiten. Der in flüssigen Grenzflächen angereicherte Stoff kann aus einer gasförmigen oder flüssigen Nachbarphase, aber auch aus dem Inneren der Adsorbensphase selbst stammen. Anreicherungen an Flüssigkeitsgrenzflächen können meist nur indirekt über Veränderungen der Oberflächen- bzw. Grenzflächenspannungen ermittelt werden.

6.6.4.1.1. *Thermodynamische Beschreibung.* Aus der Gibbsschen Fundamentalgleichung für die Grenzflächenphase (↑ 6.6.2.1.) folgt allgemein für die *Oberflächenspannung flüssiger Mischphasen* $\sigma = g^\sigma - \sum \mu_i \Gamma_i$. $\Gamma_i = n_i^\sigma / A$ ist die Stoffmenge der Komponente i in der Flächeneinheit der Grenzflächenphase, also die Über- oder Unterschußkonzentration gegenüber dem Phaseninneren; häufig wird diese Größe schlechthin als *Grenzflächenkonzentration* bezeichnet. Im Unterschied zu reinen Flüssigkeiten (↑ 6.6.2.2.1.) hängt hier die Oberflächenspannung σ noch wesentlich von den chemischen Potentialen μ_i der Komponenten der Mischphase ab.

Weiter ergibt sich aus den angeführten thermodynamischen Beziehungen die allgemeine Form der *Gibbsschen Adsorptionsgleichung* $A\,\mathrm{d}\sigma + s^\sigma\,\mathrm{d}T + \sum n_i^\sigma\,\mathrm{d}\mu_i = 0$, die für konstante Temperatur einen Zusammenhang zwischen der Änderung der Oberflächenspannung bei Änderung der chemischen Potentiale, also letztlich der Aktivitäten bzw. der Konzentrationen, liefert: $\mathrm{d}\sigma = -\sum \Gamma_i\,\mathrm{d}\mu_i$. Bei einer verdünnten Lösung mit der Konzentration c_B des gelösten Stoffs B folgt mit der Gleichung $\mathrm{d}\mu_B = RT\,\mathrm{d}\ln c_B$ schließlich für die Grenzflächenkonzentration Γ_B des adsorbierten Stoffs B: $\Gamma_B = -\frac{c_B}{RT}\left(\frac{\mathrm{d}\sigma}{\mathrm{d}c_B}\right)_T$. R ist die molare Gaskonstante und T die Temperatur. Änderungen der Oberflächenspannung mit der Konzentration lassen sich bei flüssigen Mischphasen experimentell bestimmen.

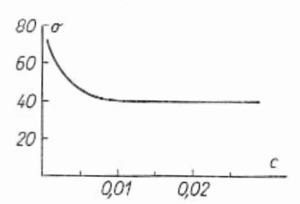


Abb. 6.6.4-1 Oberflächenspannungen σ (in $\mathrm{m\,N\,m^{-1}}$ wäßriger Lösungen von Natriumdodecylsulfat $CH_3\text{-}(CH_2)_n\text{-}O\text{-}SO_3Na$ bei 20 °C; c in $\mathrm{mol\,l^{-1}}$

6.6.4.1.2. *Grenzflächenaktivität.* Vor allem in Wasser bewirken viele gelöste Stoffe mit steigender Konzentration eine meist nichtlineare *Abnahme der Oberflächenspannung*. $\mathrm{d}\sigma/\mathrm{d}c_B < 0$ bedeutet nach der Gibbsschen Adsorptionsgleichung $\Gamma_B > 0$, es findet demnach eine Anreicherung des gelösten Stoffs an der Phasengrenze statt. Je größer c_B ist, um so größer wird der Konzentrationsunterschied zwischen Phasengrenzschicht und Flüssigkeitsinnerem.

Von besonderer Bedeutung sind Verbindungen, die in einer Lösung schon bei *sehr kleinen Konzentrationen* die Oberflächenspannung gegenüber dem reinen Lösungsmittel erheblich erniedrigen (Abb. 6.6.4-1). Sie werden als *grenzflächenaktive* (kapillaraktive) *Stoffe* oder *Tenside* bezeichnet. Für Wasser sind dies z. B. viele Stoffe, deren Moleküle aus einem unpolaren („hydrophoben") und einem polaren

Natriumdodecylsulfat
$CH_3-(CH_2)_{11}-O-SO_3Na$ Na^+

hydrophober — (hydrophiler) Molekülteil

K^+

a) *Kaliumoleat* $CH_3-(CH_2)_7-CH=CH-(CH_2)_7-COOK$

b) *Trimethylhexylphenolpolyglykolether (1:9)*

Abb. 6.6.4-2 Beispiele für den Molekülaufbau wichtiger Gruppen von Tensiden; a) anionaktive Tenside, b) nichtionogenes Tensid

(„hydrophilen") Molekülteil bestehen. Der unpolare Teil ist meist eine Kohlenwasserstoffkette oder ein anderer größerer Kohlenwasserstoffrest; am stärksten hydrophob sind perfluorierte Molekülteile. Die polaren Gruppen von Tensidmolekülen, z. B. $-COOH$, $-OSO_3H$, $-SO_3H$, bewirken infolge ihrer Fähigkeit, zu Wassermolekülen H-Brücken ausbilden zu können, eine gewisse Löslichkeit im Wasser. Da aber die Anziehungskräfte der Wassermoleküle untereinander stärker sind als die zu den organischen Molekülen, reichern sich letztere bevorzugt und orientiert an der Phasengrenze an. Ab einer bestimmten Konzentration des Gelösten bleibt die Oberflächenspannung der Lösung oftmals praktisch konstant, da dann die Oberfläche mit einer geschlossenen Schicht des gelösten Stoffs bedeckt ist.

Typische Beispiele für den Aufbau von Tensiden sind in der Abb. 6.6.4-2 dargestellt. Wegen ihrer polaren Gruppen zerfallen die meisten grenzflächenaktiven Verbindungen in wäßriger Lösung in Ionen. Verbindungen, bei denen der hydrophobe Molekülteil negativ geladen ist, werden als *anionaktive Tenside* bezeichnet. Von geringerer praktischer Bedeutung sind die *kationaktiven Tenside*, bei denen der hydrophobe Molekülteil positiv geladen vorliegt. Eine steigende Bedeutung haben als dritte Gruppe die *nichtionogenen Tenside*, die in wäßriger Lösung neutral sind und bei denen die hydrophile Molekülfunktion von einer Anhäufung schwächer polarer Gruppen, z. B. von Ethergruppen —C—O—C—, in einem größeren Molekül herrührt.

Tenside liegen nur in stark verdünnten Lösungen als Einzelmoleküle vor. Ab einer bestimmten, für jedes Tensid charakteristischen Konzentration vereinigen sich die Moleküle in der Lösung zu größeren Aggregaten, den *Mizellen.* Diese haben die Größe und Eigenschaften von Kolloiden (↑ 6.6.7.2.), so daß die Lösungen oberhalb der *kritischen Mizellbildungskonzentration* die Eigenschaften kolloiddisperser flüssiger Systeme haben. Bei dieser Konzentration, die um so niedriger liegt, je größer der hydrophobe Molekülteil ist, treten sprunghafte Änderungen vieler Eigenschaften der Tensidlösungen auf.

Die Mizellen bestehen aus größenordnungsmäßig 20000 bis 30000 Einzelmolekülen, von denen bei anion- oder kationaktiven Tensiden ein bestimmter Anteil ionisiert ist, so daß die Molekülaggregate als Ganzes elektrisch geladen sind. Sie treten außerdem in geordneten Strukturen in block-, band- oder stäbchenartiger sowie in Kugelform mit schalenartigem Aufbau

auf; in jedem Fall sind in wäßrigen Lösungen die Einzelmoleküle in der Mizelle so angeordnet, daß ihre hydrophoben Molekülteile in das Mizellinnere gerichtet sind, während die polaren Gruppen an der „Oberfläche" der Mizelle den Kontakt zum Wasser bewirken.

Die verbreitetste Anwendung finden höherkonzentrierte Tensidlösungen wegen der Fähigkeit der Mizellen, in Wasser wenig oder unlösliche flüssige Stoffe oder feste Partikel in oft beträchtlichem Umfang in sich aufnehmen zu können, wodurch diese in die flüssige Phase überführt werden. Diese Erscheinung wird als *Solubilisation* bezeichnet und stellt einen wesentlichen Teil jedes Wasch- oder Reinigungsvorgangs bei Verwendung grenzflächenaktiver Stoffe dar.

6.6.4.1.3. *Negative Adsorption.* Die Gibbssche Adsorptionsgleichung erfaßt auch den Fall, daß mit steigender Konzentration die Oberflächenspannung der Lösung zunimmt: $d\sigma/dc_B > 0$. $\Gamma_B < 0$ bedeutet, daß die gelösten Teilchen in der Phasengrenzschicht in niedrigerer Konzentration vorliegen als im Flüssigkeitsinneren und dieser „negative" Konzentrationsunterschied mit steigender Konzentration größer wird. Man spricht daher von einer *negativen Adsorption* oder Abreicherung. Molekulartheoretisch ist sie dadurch zu erklären, daß die Anziehungskräfte zwischen den Teilchen des Gelösten und des Lösungsmittels stärker sind als zwischen den Lösungsmittelmolekülen untereinander.

Man bezeichnet alle Verbindungen mit einem solchen Verhalten als *grenzflächeninaktive* (kapillarinaktive) *Stoffe*. Für Wasser als Lösungsmittel sind dies insbesondere viele anorganische Elektrolyte (Tab. 6.6.4-3).

6.6.4.2. Adsorption an Grenzflächen fester Stoffe. Hier ist in erster Linie die Adsorption von Teilchen aus gasförmigen oder flüssigen Nachbarphasen von Interesse. Die adsorbierten Mengen können aus Messungen der Partialdruck- bzw. Konzentrationsänderungen des Adsorptivs in der fluiden Phase ermittelt werden. In neuerer Zeit werden verstärkt auch andere Methoden, z. B. die IR-Spektroskopie, herangezogen, die zusätzliche Informationen, etwa über Strukturveränderungen in den adsorbierten Teilchen infolge der Wechselwirkungen mit den Oberflächenbausteinen des Adsorbens oder über die Bindungskräfte in den Adsorptionskomplexen, zu gewinnen gestatten.

6.6.4.2.1. *Adsorption aus reinen Gasphasen.* Die Grenzflächenkonzentration Γ hängt primär von der stofflichen Natur der Partner, darüber hinaus vom Druck p des Adsorptivs in der Gasphase und von der Temperatur T ab.

Im Bereich niedriger Drücke gelten bei konstanter Temperatur lineare Zusammenhänge der Form $\Gamma = \text{const} \cdot p$. Bei höheren Drücken wird häufig ein Verlauf entsprechend Abb.

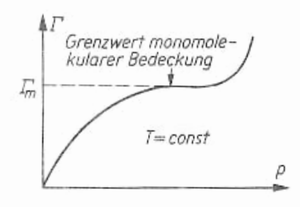


Abb. 6.6.4-4 Adsorptionsisotherme eines Gases an einer festen Oberfläche (schematisch); Γ Grenzflächenkonzentration, p Druck

Tab. 6.6.4-3 Oberflächenspannungen σ (in mNm^{-1}) wäßriger Elektrolytlösungen in Abhängigkeit von der Konzentration c (in Masse-%) bei 20 °C ($\sigma_W = 72{,}82\ mNm^{-1}$)

Salz	c/σ				
NaCl	0,58 / 72,95	2,84 / 73,75	5,43 / 74,39	14,92 / 77,65	22,62 / 80,95
KCl	0,74 / 72,99	3,60 / 73,45	6,93 / 74,15	13,88 / 75,55	22,97 / 78,28
NaBr			4,98 / 73,45	13,37 / 74,75	23,00 / 76,55
$NaNO_3$	0,85 / 72,87		4,08 / 73,75	14,53 / 75,15	29,82 / 78,35
Na_2CO_3		2,58 / 73,45	5,03 / 74,05	13,72 / 76,75	
$MgCl_2$	0,94 / 73,07		4,55 / 74,00	16,00 / 79,15	22,30 / 82,95

6.6.4-4 gefunden: Bei bestimmten Drücken streben die *Adsorptionsisothermen* zunächst einem Grenzwert zu, der offenbar dann erreicht ist, wenn alle verfügbaren Adsorptionsplätze auf der Adsorbensoberfläche mit Adsorbatteilchen besetzt sind, d. h. eine *monomolekulare Adsorptionsschicht* gebildet ist. Bis dorthin bezeichnet man diesen Verlauf auch als *Langmuirisotherme*, da er in vielen Fällen durch eine 1918 von LANGMUIR aus kinetischen Betrachtungen abgeleitete Beziehung $\Gamma = \Gamma_m bp/(1 + bp)$ beschrieben werden kann; b ist eine vom speziellen System abhängige Konstante; Γ_m ist die Grenzflächenkonzentration bei gerade vollständiger monomolekularer Bedekkung. Häufig recht gut erfüllt ist aber auch die nach FREUNDLICH benannte empirische Gleichung $\Gamma = k'p^{1/m}$ mit den Konstanten k' und m. Oftmals steigt bei weiterer Erhöhung des Drucks die adsorbierte Menge erneut stark an (↑ Abb. 6.6.4-4). Hierbei geht die monomolekulare in eine *multimolekulare Bedeckung* über, indem auf der ersten Adsorbatschicht weitere Schichten gebunden werden. Für die Beschreibung und Deutung dieser *Mehrschichtadsorptionen* eignet sich insbesondere die von BRUNAUER, EMMETT und TELLER als Erweiterung der Langmuirisotherme aufgestellte, sogenannte *BET-Gleichung*

$$\Gamma = \Gamma_m cp/(p_s - p)\,[1 + (c - 1)\,p/p_s].$$

p_s ist der Sättigungsdampfdruck des Adsorbats bei der betreffenden Temperatur und c eine kompliziert zusammengesetzte Konstante. Die BET-Gleichung wird vor allem zur Ermittlung der Oberflächengrößen fester Adsorbentien benutzt. Dazu wird neben Γ_m der Flächenbedarf eines adsorbierten Atoms oder Moleküls benötigt. Solche Messungen haben u. a. ergeben, daß die tatsächlichen Oberflächengrößen vieler pulverförmiger oder körniger fester Stoffe wie etwa Aktivkohle, Tonerden, Aluminiumoxid, Kieselsäure oder Silicagel viele hundert bis über tausend m^2g^{-1} betragen können, weil deren einzelne Körner von einer Vielzahl feinster Poren durchzogen sind.

6.6.4.2.2. Bei der *Adsorption aus fluiden Mischphasen* werden meist mehrere Bestandteile gleichzeitig, wenn auch in unterschiedlichem Umfang, an der festen Oberfläche angereichert. Dabei kann es auch zu einer *Adsorptionsverdrängung* kommen, indem ursprünglich adsorbierte Teilchen durch andere ersetzt werden, wenn letztere stärkere Wechselwirkungen zu den Oberflächenbausteinen ausüben können.

Speziell bei Adsorptionen aus wäßrigen Mischphasen sind für die Anreicherung an festen Oberflächen zwei Ursachen möglich. Zum einen können die Teilchen des betreffenden Adsorptivs stärkere Wechselwirkungen zu den Bausteinen des Adsorbens als zu den Lösungsmittelmolekülen aufweisen. Bei einem Tensid würden sich dann z. B. die polaren Gruppen zum Festkörper hin und die Kohlenwasserstoffketten in Richtung der flüssigen Phase orientieren, die feste Oberfläche wird *hydrophobiert*. Sind die Wechselwirkungen zwischen den Tensid- und den Lösungsmittelmolekülen insgesamt relativ schwach, können sich andererseits die unpolaren Molekülteile an der festen Oberfläche verankern und die polaren Gruppen zur wäßrigen Phase ausrichten, die Festkörperoberfläche wird *hydrophiliert*. Ein durch geeignete Tenside und entsprechende Verfahrensbedingungen gezielt gesteuertes Zusammenspiel dieser beiden Mechanismen ist u. a. eine wesentliche Grundlage bei der als *Flotation* bezeichneten Schwimmaufbereitung von Erzen, d. h. bei der Trennung von Erz und Gangart oder von unterschiedlichen Erzen. Von der Hydrophobierung oder der Hydrophilierung fester Oberflächen, z. B. der Hydrophobierung durch Halogensilane oder Siloxane, wird auch verbreitet für andere Zwecke Gebrauch gemacht.

Bei Adsorptionen aus fluiden Mischphasen kommt es häufig zu einer *selektiven Adsorption*, indem einzelne Bestandteile bevorzugt angereichert werden. Es sind dies stets diejenigen Teilchenarten, die mit den Oberflächenbausteinen des festen Adsorbens am stärksten in Wechselwirkung treten. Die selektive Adsorption findet vielfältige praktische Nutzung, z. B. als Gasreinigungsverfahren für Industriegase (Entfernen von H_2S oder anderen Bestandteilen aus Synthese- oder Naturgasen) oder für die Atemluft mit Hilfe von Gasmasken, für die Reinigung von flüssigen Lösungen etwa von gefärbten Verunreinigungen (u. a. bei der Zuckergewinnung oder bei der Arzneimittelherstellung) sowie als Grundlage der chromatographischen Adsorptionsanalyse.

Adsorptionserscheinungen insbesondere an hochenergetischen Oberflächen werden sehr stark von den stets vorhandenen *strukturellen* und *energetischen Heterogenitäten* dieser Oberflächen bestimmt. Dabei werden zuerst immer die Stellen größter Oberflächenenergie, dann die weniger energiereichen besetzt. Umgekehrt gestatten Adsorptionsuntersuchungen Aussagen über die energetischen Heterogenitäten „reiner" oder durch vorangegangene Adsorptionen bzw. Oberflächenreaktionen modifizierter Festkörperoberflächen.

6.6.4.2.3. *Temperaturabhängigkeit der Adsorption.* Adsorptionsvorgänge führen immer zu energieärmeren Zuständen und sind daher stets exotherm. Mit steigender Temperatur T wird deshalb das Adsorptions-Desorptions-Gleichgewicht nach der Seite der Desorption verschoben, die adsorbierte Menge nimmt ab. Näherungsweise gilt bei Adsorptionen aus der

Gasphase $(\mathrm{d}\ln p/\mathrm{d}T)_\Gamma = -\Delta_{\mathrm{Ads}}H_{\mathrm{st}}/(RT^2)$. p ist der Gasdruck des Adsorptivs, dem eine bestimmte Adsorbatmenge Γ entspricht, R die molare Gaskonstante und $\Delta_{\mathrm{Ads}}H_{\mathrm{st}}$ die *isostere molare Adsorptionsenthalpie*, d. h. die bei konstanter adsorbierter Menge auftretende Adsorptionsenthalpie. Die danach aus Adsorptionsisothermen bei verschiedenen Temperaturen bestimmbaren Adsorptionsenthalpien sind oft stark druckabhängig. Bei niedrigen Bedeckungsgraden drückt sich darin die energetische Heterogenität der festen Oberflächen aus, bei höheren die Wechselwirkungen der Adsorbatteilchen untereinander.

Allgemein sind die Adsorptionsenthalpien bei der physikalischen Adsorption im Unterschied zur Chemisorption (↑ 6.6.4.3.) relativ klein; wegen des alleinigen Wirkens zwischenmolekularer Wechselwirkungskräfte im Adsorptionskomplex liegen sie in der Größenordnung der Kondensationsenthalpien des Adsorptivs, d. h. zwischen 5 und 25 $\mathrm{kJmol^{-1}}$, maximal 40 $\mathrm{kJ\,mol^{-1}}$.

6.6.4.3. Chemisorption

6.6.4.3.1. *Allgemeine Kennzeichnung.* Bei Festkörpern mit hochenergetischen Oberflächen, vor allem bei Metallen und Ionenverbindungen, bleibt der Adsorptionsvorgang häufig nicht auf der Stufe der physikalischen Adsorption stehen. Die Wirkung der starken Metallbindungs- oder Coulombkräfte auf die Adsorbatteilchen führt vielmehr dazu, daß ihre innermolekularen Bindungen stark deformiert und teilweise gelöst werden. In diesem *chemisorbierten Zustand* liegen die Teilchen des Adsorbats dann in einer Form vor, die einer *Oberflächenverbindung* nahekommt.

Zu den charakteristischen Merkmalen der Chemisorption gehören hohe *Chemisorptionsenthalpien* von etwa 80 bis 500 $\mathrm{kJmol^{-1}}$, also in der Größenordnung chemischer Reaktionsenthalpien. Im Unterschied zur physikalischen Adsorption ist die Chemisorption weiterhin durch eine ausgeprägte *Spezifität* bezüglich der stofflichen Natur der Partner charakterisiert. Deshalb kann es auch nicht zur Ausbildung multimolekularer Schichten kommen; bei vollständiger monomolekularer Bedeckung der Adsorbensoberfläche mit dem Chemisorptionskomplex ist die Chemisorption beendet. Schließlich erfordert der Übergang von der Physisorption zur Chemisorption eine merkliche Aktivierungsenergie, wodurch der Vorgang bei niedrigen Temperaturen stark gehemmt sein kann. Die starken Bindungskräfte im Chemisorptionskomplex haben andererseits eine weitgehende Irreversibilität des Vorgangs zur Folge; ein Chemisorptions-Desorptions-Gleichgewicht stellt sich daher höchstens bei hohen Temperaturen ein.

6.6.4.3.2. *Molekulartheoretische Beschreibung.* Die Energieverhältnisse beim Übergang eines Teilchens zunächst in den Physisorptionszustand und aus diesem in den Zustand der Chemisorption sind am Beispiel der Adsorption von O_2 an einer Metalloberfläche in Abb. 6.6.4-5 schematisch dargestellt.

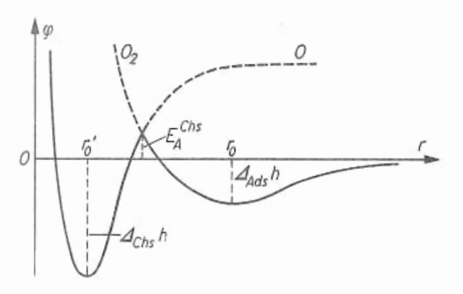


Abb. 6.6.4-5 Potentialkurven der Physisorption und der Chemisorption von O_2 an einer Metalloberfläche (schematisch); r Abstand des Adsorbatteilchens von der Adsorbensoberfläche

Bei Annäherung des O_2-Moleküls an die Oberfläche (Ordinate) wird im Abstand r_0 zunächst ein Gleichgewichtszustand erreicht, in dem das Molekül physikalisch adsorbiert vorliegt. Die Tiefe des Potentialminimums entspricht der Adsorptionsenthalpie $\Delta_{\mathrm{Ads}}h$ der Physisorption. Weitere Annäherung erfordert entsprechend dem Verlauf der Potentialkurve eine Energiezufuhr. Ist die Energie des O_2-Moleküls gleich oder größer als die Aktivierungsenergie E_A^{Chs}, dann kann es unter Deformation seiner Elektronenhülle in den Gleichgewichtszustand der Chemisorption beim Abstand r_0' übergehen. Die Tiefe dieses Potentialminimums entspricht der größeren Chemisorptionsenthalpie $\Delta_{\mathrm{Chs}}h$. Beim Übergang in den Zustand der Chemisorption wird die molekulare Struktur der adsorbierten Teilchen im Kraftfeld der Oberflächenbausteine des Adsorbens wesentlich verändert, so daß z. B. anstelle von O_2-Molekülen nur noch locker miteinander verbundene Sauerstoffatome auf der Oberfläche sitzen. Daher entspricht die Potentialkurve der Chemisorption ungefähr derjenigen einer Adsorption von O-Atomen. Mit steigender Temperatur kann eine wachsende Zahl von Molekülen die hemmende Energiebarriere E_A^{Chs} überwinden, d. h. die insgesamt chemisorbierte Menge steigt trotz Temperaturerhöhung an, bis der Vorgang bei Ausbildung einer geschlossenen Chemisorptionsschicht beendet wird.

6.6.4.3.3. *Praktische Bedeutung; reale Festkörperoberflächen.* Die bei einer Chemisorption auftretende Deformation der Elektronenhüllen der chemisorbierten Teilchen führt dazu, daß diese u. a. auch leichter mit anderen Verbindungen in Reaktion treten können. Die Chemi-

sorption ist daher eine der wesentlichen Grundlagen der heterogenen Katalyse (↑ 7.3.3.3.). Reine hochenergetische Festkörperoberflächen, besonders von Metallen, sind praktisch nur im Ultrahochvakuum und durch spezielle Verfahren, z. B. der Ionentechnologien, zu erhalten. Schon bei tiefen Temperaturen werden geringste Spuren gasförmiger Verbindungen einschließlich der permanenten Gase adsorbiert; hinreichendes Angebot an Adsorptiv führt zur Ausbildung geschlossener Chemisorptionsschichten. Bei normaler Temperatur bilden sich aus ihnen oftmals bereits echte Verbindungs-, z. B. Oxid- oder Nitridschichten.

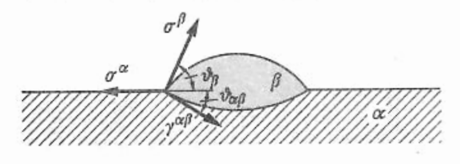


Abb. 6.6.5-1 Schwimmender Tropfen (Erläuterung im Text)

Unter ***natürlichen Bedingungen*** sind selbst auf Edelmetallen dünnste Oxidschichten vorhanden. Die technisch wichtigen Gebrauchsmetalle wie Eisen, Nickel, Zink, Aluminium oder Kupfer sind mit dickeren, oft nichtstöchiometrisch zusammengesetzten Oxidschichten bedeckt, die durch Adsorption weiteren Sauerstoffs aus der Atmosphäre und Diffusionsprozesse in der Schicht auf festkörperreaktivem Wege entstehen. So sind alle „Metall"- wie die Glas- oder Silicatoberflächen eigentlich hochenergetische *Oxidoberflächen*. An ihnen werden weiter aus der Luft die stets vorhandenen Wassermoleküle adsorbiert. Ein Teil davon reagiert chemisch mit den Oberflächenoxiden unter Bildung von Hydroxylgruppen, an denen dann weitere Wassermoleküle über H-Brücken fest gebunden werden. Es bedarf bei einer Vielzahl derartiger Oberflächen Temperaturen von weit über 100 °C und längere Zeit der Erhitzung im Vakuum, um diese Wassermoleküle vollständig wieder abzubauen. Bei Normaltemperatur und bereits kleinen relativen Luftfeuchtigkeiten von etwa 10% werden z. B. an Al-, Fe-, Cu- oder Glasoberflächen an die erste Wasserschicht bis zu 20 weitere Schichten von Wassermolekülen gebunden. Die Dicke dieser „Wasserhaut" an „trockenen" Metall- oder Silicatoberflächen nimmt mit steigender Luftfeuchtigkeit zu.

Auf Metallen, die mit der normalen Atmosphäre in Kontakt stehen, sind die Oxidschichten häufig noch durch Sulfide, Sulfate, Carbonate oder andere Salze verunreinigt, die durch Oberflächenreaktionen mit entsprechenden Luftbestandteilen bzw. -verunreinigungen entstehen. Schließlich werden an den Wasserfilmen oft noch organische Verbindungen adsorbiert, die in unsauberen Luftatmosphären ebenfalls immer vorhanden sind oder mit denen die festen Oberflächen bei der Aufbewahrung oder Behandlung in Berührung kommen. Alle diese möglichen Physisorptions-, Chemisorptions- und Reaktionsschichten wirken sich natürlich stark auf viele Verarbeitungs- oder Gebrauchseigenschaften der festen Stoffe aus.

6.6.5. Spreitung und Oberflächenfilme

Unter *Spreitung* versteht man die spontane Ausbreitung einer kleinen Menge flüssiger oder fester Stoffe auf einer flüssigen oder festen Oberfläche. Dabei entsteht in vielen Fällen ein *Oberflächenfilm* aus dem gespreiteten Stoff.

6.6.5.1. Spreitung. Bleibt ein auf eine Flüssigkeitsoberfläche aufgebrachter Tropfen einer spezifisch leichteren Flüssigkeit β (z. B. eines Alkans), die in der Flüssigkeit α (z. B. Wasser) schwer löslich ist, in Form einer Linse liegen (Abb. 6.6.5-1), dann entspricht dies wegen der vorhandenen Beweglichkeit aller Teilchen einem stabilen Gleichgewichtszustand. Die *allgemeinste Gleichgewichtsbedingung* für ein solches System ist dadurch gegeben, daß die Summe aller Oberflächen- bzw. Grenzflächenenergien $(\sigma^\alpha A^\alpha + \sigma^\beta A^\beta + \gamma^{\alpha\beta} A^{\alpha\beta})$ einen Minimalwert annimmt. Jede der zwei Oberflächenspannungen σ und die Grenzflächenspannung γ wirkt für sich allein in dem Sinne, daß sie die betreffende Grenzfläche zu verkleinern sucht. In der (kreisförmigen) *Dreiphasengrenzlinie* sind alle drei Spannungen gleichzeitig wirksam. Im Gleichgewicht halten sich ihre horizontalen Komponenten gerade die Waage, es gilt daher als *spezielle Gleichgewichtsbedingung* $\sigma^\alpha >= \sigma^\beta \cos\vartheta_\beta + \gamma^{\alpha\beta} \cos\vartheta_{\alpha\beta}$. Je größer der Unterschied zwischen σ^α und der Summe $(\sigma^\beta + \gamma^{\alpha\beta})$ ist, um so kleiner werden die Winkel ϑ_β und $\vartheta_{\alpha\beta}$, um so breiter fließt der Tropfen auseinander. Bei Flüssigkeiten α mit großer Oberflächenspannung, z. B. Wasser oder Hg, tritt sogar häufig der Fall auf, daß $\sigma^\alpha > (\sigma^\beta + \gamma^{\alpha\beta})$; dann ist die Tropfenform nicht mehr stabil und die Flüssigkeit β *spreitet* auf α. Für Wasser trifft dies z. B. auf alle gesättigten und ungesättigten cyclischen Kohlenwasserstoffe oder die polaren Öle und Siliconöle zu; auf Hg spreiten mit Ausnahme von Wasser praktisch alle Flüssigkeiten.

Als Kenngröße für die Spreitung verwendet man die nach $S^{\beta\alpha} = \sigma^\alpha - (\sigma^\beta + \gamma^{\alpha\beta})$ definierte *Spreitungsspannung* $S^{\beta\alpha}$, die auch als *Spreitungskoeffizient* bezeichnet wird. Nach der *Spreitungsbedingung* $S^{\beta\alpha} \geqq 0$ ist ein Spreiten nur zu erwarten, wenn eine Flüssigkeit mit niedrigerer Oberflächenspannung auf eine mit größerer gebracht wird.

Spreitung tritt nur ein, wenn die Wechselwirkungen zwischen den Teilchen der Flüssigkeiten α und β stärker sind als diejenigen zwischen den Teilchen der Flüssigkeit β untereinander. Ein Spreiten vor allem auf Wasser kann daher auch aus dem festen Zustand heraus erfolgen, wenn die betreffenden Stoffe einen relativ niedrigen Schmelzpunkt besitzen und aus Molekülen mit größeren hydrophoben Molekülteilen und stark polaren Gruppen bestehen. Beispiele sind die höheren Fettsäuren und ihre Erdalkalisalze sowie Kampfer. Diese und andere feste Stoffe können aber auch dadurch in einen gespreiteten Zustand überführt werden, daß man sie in einem leichtflüchtigen Lösungsmittel löst und die Lösung auf die Flüssigkeitsoberfläche aufbringt, auf der sie spreitet, wobei das Lösungsmittel verdunstet.

Es gibt viele organische Verbindungen mit polaren Gruppen, die beim Auftropfen z. B. auf eine Wasseroberfläche im ersten Moment spreiten, weil die polaren Gruppen mit den Wassermolekülen in starke Wechselwirkungen treten, sich dann aber sehr schnell wieder zu einer flachen Linse zusammenziehen. Die *Anfangsspreitungsspannung* muß also offenbar positiv sein, die *Endspreitungsspannung* dagegen einen negativen Wert haben. Die Wasseroberfläche ist dann neben der Linse mit einem monomolekularen Film des gespreiteten Stoffs bedeckt, in dem die Teilchen mit ihren polaren Enden zu den Wassermolekülen, mit ihren unpolaren Molekülteilen dagegen nach der Gasphase orientiert sind. Dadurch wird die Oberflächenspannung der Unterlage, auf der die Linse ruht, stark herabgesetzt. Man bezeichnet Flüssigkeiten, die auf ihrer eigenen Monoschicht nicht mehr spreiten können, als *autophobe Flüssigkeiten.*

6.6.5.2. Spreitungsgeschwindigkeit und Oberflächendruck. Polare Öle spreiten auf Wasser mit mittleren Geschwindigkeiten von etwa 10 bis 20 cms^{-1}. Die *Spreitungsgeschwindigkeit* hängt in komplizierter Weise von der Spreitungsspannung, den Viskositäten beider Flüssigkeiten, der Temperatur und anderen Faktoren, z. B. der bereits erfolgten Ausbreitung, ab.

Der sich ausbreitende *Oberflächenfilm* übt an seiner Peripherie einen meßbaren „Schub" aus, z. B. auf eine auf der Wasseroberfläche schwimmende Barriere. Dieser ist physikalisch gleich der Kraft je Längeneinheit, die aufgewandt werden muß, damit er die jeweils eingenommene Fläche beibehält. Nach erfolgter Ausbreitung ist dieser „Druck" numerisch gleich dem Unterschied zwischen der Oberflächenspannung der reinen Flüssigkeitsoberfläche, σ^α, und der mit einem Oberflächenfilm bedeckten, σ^α_F, d. h. der Oberflächenspannungserniedrigung durch die Veränderung der stofflichen Zusammensetzung der Flüssigkeitsoberfläche. Diese Differenz bezeichnet man daher als *Oberflächendruck* (*Filmdruck, Spreitungsdruck*) π, für den gilt $\pi = \sigma^\alpha - \sigma^\alpha_F$. Der Oberflächendruck ist eine sehr zweckmäßige Kenngröße zur Beschreibung von Veränderungen des grenzflächenenergetischen Zustands beliebiger Grenzflächen mit der stofflichen Zusammensetzung. Oberflächenfilme können ja nicht nur durch Spreitung, sondern auch durch Adsorption von Bestandteilen angrenzender Phasen gebildet werden.

6.6.5.3. Oberflächenfilme auf Flüssigkeiten. Diese sind, insbesondere auf Wasser, schon sehr lange bekannt. Bereits 1765 beobachtete FRANKLIN, daß sich Olivenöl auf Wasser zu einem dünnen Film von 2,5 nm Dicke ausbreitet. Aus heutiger Sicht können diese und ähnliche Beobachtungen als die ersten experimentellen Beweise der Neuzeit für den korpuskularen Charakter der Stoffe angesehen werden: Diese Filme können nicht dünner werden, da 2,5 nm die Molekülgröße ist.

Oberflächenfilme können in verschiedenen Formen auftreten. Neben den *monomolekularen Filmen* ist aber auch die Bildung *multimolekularer Filme* möglich. Solche bezeichnet man als *Duplexfilme*, wenn sie so dick sind, daß sie noch die Volumeneigenschaften des betreffenden Stoffs und an der Phasengrenze zur Gasatmosphäre dessen Oberflächenspannung aufweisen. Innerhalb der monomolekularen Filme können die Moleküle je nach der Stärke ihrer Wechselwirkungskräfte untereinander und zu den Wassermolekülen mehr oder weniger frei beweglich sein (*gasanaloge Filme*) oder einen flüssigkeits- bzw. sogar fettkörperähnlichen geschlossenen Film bilden (*kondensierte Filme*). In letzteren liegt ein hoher Ordnungsgrad vor. Bei Filmen aus längerkettigen Kohlenwasserstoffresten und stark polaren Endgruppen ergibt sich unabhängig von der Kettenlänge oftmals der gleiche Flächenbedarf von rund 2 nm^2 pro Molekül, weil die Teilchen mit den polaren Gruppen in der Wasseroberfläche „verankert" sind und die Kohlenwasserstoffketten senkrecht und parallel zueinander in die Gasphase ragen.

Jeder Oberflächenfilm bewirkt wesentliche Veränderungen aller Grenzflächeneigenschaften der filmtragenden Flüssigkeit, die vielfältige praktische Auswirkungen haben. Insbesondere durch kondensierte Filme wird die Verdampfung der Trägerflüssigkeit merklich herabgesetzt, da hierzu ein Massetransport durch den Film erforderlich ist. In tropischen Ländern vermindert man damit Verdunstungsverluste bei offenen Wasserreservoiren. Bekannt sind auch Veränderungen der Lichtreflexion (Auftreten von Interferenzfarben von Ölfilmen auf Wasserpfützen).

6.6.5.4. Oberflächenfilme auf festen Oberflächen. Spreitung von Flüssigkeiten tritt auch auf hochenergetischen festen Oberflächen auf. Von größerer praktischer Bedeutung sind hier jedoch solche Filme, die durch Adsorption aus fluiden Nachbarphasen gebildet werden, z. B. die multimolekularen Wasserfilme auf Metalloberflächen (↑ 6.6.4.3.3.). Durch die meist vorhandenen Oberflächenheterogenitäten treten dabei oft *Inselstrukturen* auf: An bevorzugten Oberflächenstellen wird zunächst eine mehr oder weniger große Anzahl von Teilchen in Pakkungen übereinander adsorbiert. Erst bei weiterem Angebot an Adsorptivmolekülen wachsen die Inseln unter Ausbildung eines geschlossenen, dann sofort multimolekularen Films zusammen. Definierte monomolekulare Oberflächenfilme können auf weitgehend glatten, ebenen Festkörperflächen nach der Methode der *Aufbaufilme* oder mit autophoben Flüssigkeiten erzeugt werden.

Verdunstet nach dem Spreiten einer flüssigen Mischphase an einigen Stellen der Oberflächen eine Komponente, so liegt dort eine andere Oberflächenspannung vor als in der Umgebung. Das System versucht dann, die Differenz auszugleichen, indem aus der Umgebung weitere Flüssigkeit herangezogen wird. Die Erscheinung, daß durch einen Oberflächenspannungsgradienten ein Massetransport verursacht wird, bezeichnet man als *Marangonieffekt*. Am bekanntesten ist die Bildung von Flüssigkeitstropfen aus gespreiteten Filmen an senkrechten Oberflächen, z. B. das „Tränen“ von Likör oder schweren Weinen in Gläsern.

6.6.6. Benetzung

Darunter versteht man den mehr oder weniger vollständigen Kontakt von Flüssigkeiten mit festen Oberflächen, bei dem die stoffliche Natur der Phasen im wesentlichen unverändert bleibt. Als Folge eines solchen Kontakts kann es jedoch zu einem Quellen oder zu einer Auflösung des festen Stoffs, zu einer Anreicherung bestimmter Komponenten in der Grenzfläche fest–flüssig ($s - l$) durch Physisorption oder gar zu chemischen Grenzflächenreaktionen kommen.

6.6.6.1. Benetzungsgleichgewicht

6.6.6.1.1. *Thermodynamische Beschreibung.* Wird ein Tropfen einer Flüssigkeit (l) in einer inerten Gasatmosphäre (g) auf eine horizontale feste Oberfläche (s) aufgebracht, kann dieser entweder auf ihr spreiten (↑ 6.6.5.) oder die Form eines Kugelabschnitts von flacher Linse bis nahezu vollständiger Kugel annehmen (Abb. 6.6.6-1). Die *allgemeinste Gleichgewichtsbedingung* für ein solches System läßt sich wieder als Minimalwert der Summe aller Oberflächen- bzw. Grenzflächenenergien ($\sigma^s A^s + \sigma^l A^l + \gamma^{s,l} A^{s,l}$) angeben. Durch Anwendung der Gibbsschen Fundamentalgleichung für die Grenzflächenphase (↑ 6.6.2.1.) kann die *spezielle Gleichgewichtsbedingung* abgeleitet werden, wobei folgende Zusatzbedingungen berücksichtigt werden müssen: Thermodynamisches Gleichgewicht in derartigen Systemen bedeutet, daß sich auch das *Verdampfungsgleichgewicht* der Flüssigkeit und an der Phasengrenze $s - g$ das *Adsorptionsgleichgewicht* eingestellt haben muß. Im *Benetzungsgleichgewicht* ist deshalb nicht die Größe σ^s, sondern die Oberflächenspannung des mit einer Adsorptionsschicht bedeckten Festkörpers $\sigma^{s,\,ads\,l}$ wirksam. Schließlich muß die feste Oberfläche als völlig eben, strukturell und energetisch homogen und nichtdeformierbar angesehen werden, und der feste Stoff darf durch die Flüssigkeit nicht angequollen oder aufgelöst werden. Dann erhält man $\sigma^{s,\,ads\,l} = \gamma^{s,l} + \sigma^l \cos\vartheta$. Phänomenologisch läßt sich diese 1805 in vereinfachter Form empirisch gefundene *Youngsche Gleichung* leicht verstehen (↑ Abb. 6.6.6-1). In der (kreisförmigen) *Dreiphasengrenzlinie* müssen sich die horizontalen Komponenten der drei Oberflächen- bzw. Grenzflächenspannungen gerade aufheben.

Die das Benetzungsgleichgewicht charakterisierende makroskopische Größe ist der Winkel ϑ zwischen Festkörper- und Flüssigkeitsoberfläche am Ort der Dreiphasengrenzlinie. Er heißt *Randwinkel* oder *Kontaktwinkel* (exakter: Gleichgewichtsrandwinkel) und kann experimentell bestimmt werden. Die Youngsche Gleichung gilt für alle Randwinkel zwischen 0 und 180 Grad.

6.6.6.1.2. *Weitere Zusammenhänge.* Von den 4 Größen in der Youngschen Gleichung sind nur σ^l und ϑ experimentell zugänglich. Damit ist empirisch auch nur die Größe $\sigma^{s,\,ads\,l} - \gamma^{s,l} = \sigma^l \cos\vartheta = B^{s,l}$ bestimmbar, die als *Benetzungsspannung*, auch *Haft-* oder *Adhäsionsspannung*, $B^{s,l}$, bezeichnet wird. Aus ihr sind keine Aussagen über die Einzelwerte von $\sigma^{s,\,ads\,l}$ oder $\gamma^{s,l}$ möglich.

Das Benetzungsgleichgewicht läßt sich auch noch in einer anderen, für praktische Zwecke häufig gebrauchten Form darstellen. Dazu wird der Zusammenhang mit der Adhäsionsarbeit (↑ 6.6.3.1.3.) herangezogen, der hier die spezielle Form $w^{s,l}_{Ad} = \sigma^{s,\,ads\,l} + \sigma^l - \gamma^{s,l}$ annimmt. Man erhält damit die *Young-Duprèsche Gleichung* $w^{s,l}_{Ad} = \sigma^l(1 + \cos\vartheta)$. Von der allge-

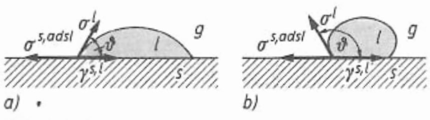


Abb. 6.6.6-1 Liegende Tropfen und Randwinkel ϑ (Erläuterung im Text)

meingültigen Youngschen Gleichung kann schließlich ein Zusammenhang mit der Oberflächenspannung σ^s der „reinen" festen Oberfläche hergestellt werden. Es ergibt sich $\sigma^s = \gamma^{s,l} + \sigma^l \cos\vartheta + \pi$, wobei π der Oberflächendruck ist (↑ 6.6.5.2.). Mit einer Adsorption von Dampf der Flüssigkeit auf der festen Oberfläche muß immer gerechnet werden, wenn die Flüssigkeit eine niedrigere Oberflächenspannung besitzt als der feste Stoff und hinreichend flüchtig ist. Nur bei festen Stoffen mit niederenergetischen Oberflächen kann der Fall eintreten, daß $\pi = 0$ oder vernachlässigbar klein ist.

6.6.6.2. Benetzungsverhältnisse in realen Systemen

6.6.6.2.1. *Benetzung und Nichtbenetzung.* Ist der in einem System $s - l - g$ auftretende Randwinkel $\vartheta < 90°$, so spricht man von einer (*unvollständigen*) *Benetzung* der festen Oberfläche durch die Flüssigkeit bzw. von einer die Oberfläche benetzenden Flüssigkeit. Nach der Youngschen Gleichung müssen für diesen Fall $\sigma^{s,\,\mathrm{ads}\,l} > \gamma^{s,l}$ und $B^{s,l}$ positiv sein. Glas oder Metalle werden von vielen Flüssigkeiten benetzt, wenn ihre Oberflächen sehr rein sind. Oftmals sind die dabei auftretenden Randwinkel sogar Null. Diesen *Grenzfall* bezeichnet man als *vollständige Benetzung*. Die Flüssigkeit spreitet dann auf der festen Oberfläche, indem eine positive Spreitungsspannung $S^{l,s} = \sigma^{s,\mathrm{ads}\,l} - (\sigma^l + \gamma^{s,l}) \geqq 0$ auftritt und die Benetzungsspannung $B^{s,l}$ den Wert der Oberflächenspannung der Flüssigkeit erreicht oder überschreitet.

Von *Nichtbenetzung* wird dann gesprochen, wenn der in einem System $s - l - g$ auftretende Randwinkel $\vartheta > 90°$ ist. Für diesen Fall muß $\sigma^{s,\,\mathrm{ads}\,l} < \gamma^{s,l}$ sein und die Benetzungsspannung einen negativen Wert haben. Einige Beispiele für große Randwinkel sind in Tab. 6.6.6-2 aufgeführt.

Der Randwinkel wird stets bestimmt durch das Verhältnis der Wechselwirkungskräfte zwischen den Oberflächenbausteinen des festen Stoffs und der Flüssigkeit zu denjenigen der Flüssigkeitsteilchen untereinander. Der formale Grenzfall $\vartheta = 180°$, der einer vollständigen *Nichtbenetzung* entsprechen würde, tatsächlich aber nie beobachtet werden konnte, ist auch theoretisch deswegen nicht zu erwarten, weil er überhaupt keinerlei Wechselwirkungen zwischen den Festkörper- und den Flüssigkeitsbausteinen über die Phasengrenze hinweg zur Voraussetzung hätte. Über Dispersionskräfte tritt jedoch grundsätzlich immer eine Wechselwirkung zwischen verschiedenartigen Teilchen in Phasengrenzen auf (↑ 6.6.3.1.2.).

6.6.6.2.2. *Benetzung hochenergetischer Oberflächen.* „Metall"-, Oxid- oder Glasoberflächen sollten durch alle normalen Flüssigkeiten vollständig benetzt werden. Mit völlig sauberen, hochpolierten festen Stoffen und reinsten Flüssigkeiten in inerten Gasatmosphären, die absolut frei von organischen Verunreinigungen und Wasserdampf waren, ist dies auch experimentell nachgewiesen. Schon ein kurzzeitiger Kontakt mit einer Gasatmosphäre von unter 1% relativer Luftfeuchtigkeit verändert jedoch den Oberflächencharakter dieser Festkörper grundlegend. Der dann vorhandene Wasserfilm verursacht eine unvollständige Benetzung für viele Flüssigkeiten, wenn deren Moleküle nicht durch einen speziellen Aufbau in der Lage sind, die adsorbierte Wasserschicht zu verdrängen. Dickere Wasserfilme (↑ 6.6.4.3.3.) überdecken die spezielle stoffliche Natur oder unterschiedliche Strukturen der Festkörperrandschichten bei der Benetzung mit vielen organischen Flüssigkeiten praktisch vollständig.

Eine analoge, lediglich noch vielfältigere und meist unkontrollierbare Wirkung haben organische Verunreinigungen auf den Festkörperoberflächen. Unter normalen Umständen verhalten sich daher die meisten Festkörper mit an sich hochenergetischen Oberflächen hinsichtlich ihrer Benetzungseigenschaften wie Stoffe mit undefinierten niederenergetischen Oberflächen. Es bedarf dann stets empirisch zu erprobender Behandlungen mit geeigneten „Reinigungsmitteln", um z. B. solche Oberflächen für bestimmte Flüssigkeiten gut benetzbar zu machen.

6.6.6.2.3. *Praktische Bedeutung.* Benetzungserscheinungen spielen in der Natur und Technik eine große Rolle. Sie beeinflussen z. B. heterogene Keimbildungsvorgänge bei der Kondensation von Dämpfen, das Eindringen von Flüssigkeiten in kapillare Hohlräume, den Wasserhaushalt in lebenden Pflanzen oder metallurgische Prozesse. Sie sind wesentliche Grundlage bei allen Wasch- und Reinigungsverfahren, bei allen Verfahren zur Beschichtung fester Stoffe mit flüssigen Stoffen oder in allen Druckprozessen bei der Farbübertragung von der Druckform auf den Bedruckstoff.

Weil die Wechselwirkungen zwischen den Oberflächenbausteinen des festen Stoffs und

Tab. 6.6.6-2 Beispiele für Nichtbenetzung; Randwinkel ϑ in Grad

System	ϑ
Hg auf Polytetrafluorethylen	≈150
Hg auf Glas	140
H_2O auf Polytetrafluorethylen	108
H_2O auf Paraffinen	105 bis 110
H_2O auf Siliconen	105 bis 108
H_2O auf Polypropylen	95
H_2O auf Polyethylen	94

den Teilchen der Flüssigkeiten dann am stärksten sind, ist vollständige Benetzung eine wesentliche Voraussetzung für das Löten, Schweißen und Kleben. Hierbei wendet man Zusätze an, die die feste Oberfläche nicht nur durch Adsorption, sondern auf chemisch-reaktivem Wege verändern; sie lösen z. B. Oxidschichten auf, wodurch das Lot in metallischen Kontakt zur Unterlage treten kann, vermitteln eine möglichst gute Haftung, z. B. von Klebstoffen, Lacken oder bei der Herstellung von Verbundwerkstoffen wie glasfaserverstärkten Polyesterharzen. Eine gleiche Zielstellung liegt dem Einsatz von bestimmten „Additivs" zu allen Schmierstoffen zugrunde. Durch die Ausbildung von durch Öl gut benetzbaren, sich selbst regenerierenden Adsorptions- und Reaktionsschichten, vor allem aus Sulfiden, kann die Reibung und damit der Verschleiß mechanisch bewegter Teile in Maschinen, Anlagen oder Motoren merklich herabgesetzt werden.

Die in der Praxis vielfältig auftretende Aufgabe, Festkörperoberflächen durch Wasser möglichst wenig benetzbar zu machen, kann im Prinzip immer dadurch gelöst werden, daß man auf die Oberfläche solche Adsorptions-, besser Reaktionsschichten aufbringt, die ihnen einen niederenergetischen Charakter verleihen. Oftmals genügt dann dazu ein kondensierter monomolekularer Film, dessen starr orientierte Moleküle mit stark hydrophoben Molekülteilen, etwa CH_3- oder CF_3-Gruppen, in die Gasphase gerichtet sind (↑ 6.6.4.2.2.).

6.6.6.3. Benetzung niederenergetischer Oberflächen

6.6.6.3.1. *Kritische Oberflächenspannung der Benetzung.* Bei Feststoffen mit niederenergetischen Oberflächen nehmen die $\cos\vartheta$-Werte mit abnehmender Flüssigkeitsoberflächenspannung σ^l linear entsprechend $\cos\vartheta = a - b\sigma^l$ zu (Abb. 6.6.6-3); a und b sind stoffabhängige Konstanten. Die $\cos\vartheta - \sigma^l$-Gerade kann stets auf den Ordinatenwert $\cos\vartheta = 1$ ($\vartheta = 0°$) extrapoliert werden. Der zugehörige Abszissenwert wird als *kritische Oberflächenspannung der Benetzung* γ_c bezeichnet. Der γ_c-Wert gibt diejenige Oberflächenspannung an, die eine Flüssigkeit höchstens besitzen darf, wenn sie auf der jeweiligen Oberfläche spreiten soll. Alle Flüssigkeiten, deren Oberflächenspannungen größer als γ_c sind, bilden auf der Oberfläche Tropfen mit Randwinkeln $> 0°$.

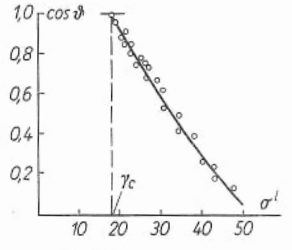


Abb. 6.6.6-3 Zismangerade zur Ermittlung der kritischen Oberflächenspannung der Benetzung γ_c in m Nm^{-1} von Polytetrafluorethylen; σ^l in m Nm^{-1}

Die γ_c-Werte sind keine fundamentalen Größen zur absoluten Charakterisierung niederenergetischer fester Oberflächen. Auf der Grundlage molekulartheoretischer Ansätze für die Grenzflächenspannungen $\gamma^{s,l}$ und für die Beiträge der verschiedenen Wechselwirkungsanteile (↑ 6.6.3.1.2.) konnte jedoch nachgewiesen werden, daß die kritischen Oberflächenspannungen γ_c mit den Oberflächenspannungen σ^s der festen Stoffe identisch sind, wenn zur Bestimmung von γ_c eine Reihe solcher Flüssigkeiten verwendet wird, die mit der festen Oberfläche *reguläre Phasengrenzen* bilden. Als regulär wird eine Phasengrenze zwischen Stoffen A und B bezeichnet, deren Moleküle jeweils untereinander die gleichen Arten von Wechselwirkungskräften ausüben und bei denen die Wechselwirkungsenergien $\varphi(r)_{AB}$ zwischen den Teilchen A und B, also über die Phasengrenze hinweg, gleich dem geometrischen Mittel der Wechselwirkungen $\varphi(r)_{AA}$ und $\varphi(r)_{BB}$ sind, $\varphi(r)_{AB} = [\varphi(r)_{AA}\varphi(r)_{BB}]^{1/2}$. Müssen zur Bestimmung von γ_c-Werten dagegen solche Flüssigkeitsreihen verwendet werden, die keine regulären Phasengrenzen mit dem Feststoff bilden, dann sind die γ_c-Werte nur Näherungswerte für σ^s.

6.6.6.3.2. *Chemische Oberflächenkonstitution.* Die Benetzungseigenschaften niederenergetischer wie auch mit entsprechenden Adsorptionsschichten bedeckter hochenergetischer fe-

Tab. 6.6.6-4 Kritische Oberflächenspannungen der Benetzung γ_c (in mNm^{-1}) fester Oberflächen aus bestimmten Atomgruppen

Atomgruppe	γ_c	Stoffbeispiel
$—CF_3$	6	Aufbaufilm aus perfluorierter Carbonsäure
$—CF_2—$	18,5	Polytetrafluorethylen
$—CF_2—CH_2—$	25	Polyvinylidenfluorid
$—CH_3$	22 bis 24	Paraffine
$—CH_3$	24	Polydimethylsiloxan
$—CH_2—$	26 bis 28	Polyethylen
$—CH_2—CH=$	33	Polystyren
$—CH_2—CHOH—$	37	Polyvinylalkohol
$—CH_2—CHCl—$	39	Polyvinylchlorid
$—CO—NH—$	41	Polyamid
$—CH_2—CHCN—$	44	Polyacrylnitril

ster Oberflächen werden weitgehend durch die in ihnen liegenden Atomgruppen bestimmt. Damit hängen auch die γ_c-Werte nahezu ausschließlich von der chemischen Natur und dem strukturellen Aufbau dieser Gruppen in der Oberfläche ab. Untersuchungen an Modellsystemen mit genau bekannten Oberflächenzusammensetzungen ergaben für typische Gruppen organischer hoch- oder auch niedermolekularer Stoffe z. B. die in Tab. 6.6.6-4 aufgeführten Werte.

Die umfangreichen Tabellen mit γ_c-Werten, auch für Atomgruppen mit weiteren Elementen, sind vielfältig praktisch nutzbar, z. B. zur Ermittlung oder zur Kontrolle von Oberflächenzusammensetzungen organischer Materialien oder Schichten durch Benetzungsuntersuchungen.

6.6.7. Disperse Systeme

In Natur und Technik gibt es zahllose *Mehrkomponentensysteme*, in denen ein oder mehrere Bestandteile, die *dispersen Bestandteile*, in „gröberer" Form in einem homogenen Medium, dem *Dispersionsmittel*, verteilt sind. Die makroskopischen Eigenschaften solcher *dispersen Systeme* werden in hohem Maße dadurch bestimmt, daß sie durch die feine Verteilung der dispersen Bestandteile z. T. sehr große „innere" Oberflächen (↑ 6.6.1.) besitzen und damit stets energiereicher sind als die entsprechenden kompakten Phasen (↑ 6.6.2.).

6.6.7.1. Begriffe, Systematik, Beispiele. Nach der Zahl der Atome, aus denen die dispersen Einzelteilchen aufgebaut sind, unterscheidet man kolloiddisperse und grobdisperse Systeme. Bestehen die dispersen Einzelteilchen aus etwa 10^3 bis 10^9 Atomen, dann liegen *kolloiddisperse Systeme* vor. Man nennt entsprechende *Dispersionen* von festen oder flüssigen Partikeln in Gasen *Aerosole*, z. B. Nebel, Sprays, Rauche, Stäube; von Gasen in Flüssigkeiten oder festen Stoffen *Schäume*, z. B. Bimsstein oder Meerschaum; von Flüssigkeiten in flüssigen Dispersionsmitteln *Emulsionen* (genauer kolloide Emulsionen), z. B. Milch; von festen Partikeln in Flüssigkeiten *Suspensionen* (genauer kolloide Suspensionen), z. B. Tusche oder viele Farben mit Farbstoffpigmenten; und von festen Partikeln in festen Stoffen *feste Sole*, z. B. viele gefärbte Gläser oder manche gefüllte Polymere. Allgemein bezeichnet man kolloide Dispersionen in beliebigen Dispersionsmitteln als *Sole*; kolloide Dispersionen in flüssigen Dispersionsmitteln sind *Lyosole* oder *kolloide Lösungen*, in Wasser als Dispersionsmittel *Hydrosole*. Die meisten kolloiddispersen Systeme sind *polydispers*, sie enthalten disperse Bestandteile in unterschiedlichen Größen und Gestalten. Diese sind im allgemeinen lichtmikroskopisch noch nicht sichtbar, verursachen aber bei seitlicher Beleuchtung eine Lichtstreuung, den *Tyndalleffekt*. Flüssige kolloiddisperse Systeme sind je nach der Art der Kolloidteilchen optisch klar, opaleszierend oder trüb. Durch Membranen geeigneter Porengröße, z. B. Pergament oder tierische Häute, werden die Partikel in Lyosolen meist zurückgehalten; sie können auf diese Weise vom Dispersionsmittel oder von niedermolekularen Bestandteilen durch *Dialyse* oder *Ultrafiltration* getrennt werden.

Der Übergang von den kolloiddispersen zu den *grobdispersen Systemen* ist fließend. Die dispersen Einzelteilchen in letzteren bestehen aus mehr als 10^9 Atomen, sie sind lichtmikroskopisch, oft sogar mit bloßem Auge, sichtbar und mit normalen mechanischen Hilfsmitteln, z. B. durch Filtration, abtrennbar. Die Systeme sind praktisch immer polydispers. Nach den Aggregatzuständen von Dispersionsmitteln und dispersen Bestandteilen unterscheidet man wie bei den kolloiddispersen Systemen Nebel, Rauche, Schäume, Emulsionen und Suspensionen. Viele solcher natürlicher und künstlicher Systeme, z. B. die Wolken, Sand-, Erd- oder Industriestäube, Schäume von Waschflotten, Schaumstoffe, Butter, Salben, Latices, Klebstoffe, Tonaufschlämmungen, Schlämme, Pasten, Beton, viele Gesteine, enthalten gleichzeitig kolloid- und grobdisperse Bestandteile.

6.6.7.2. Kolloide Systeme. Die Bezeichnung *Kolloide* wurde von GRAHAM aus dem Griechischen [*colla*, Leim; *eides*, ähnlich] abgeleitet. Verstand man darunter ursprünglich nur solche Stoffe wie Eiweiße, Stärke, Cellulose oder Harze, ist heute bekannt, daß sich im Prinzip jeder Stoff in einen kolloiden Zustand, z. B. in Form eines Lyosols, überführen läßt. Neben seiner Bindung an das Wasser ist Leben in allen seinen Formen an die speziellen Eigenschaften dieses Zustands gebunden. Kolloide Dispersionen bestimmen zahlreiche Vorgänge im Erdboden, Verwitterungs- und Korrosionsprozesse. Etwa 10% aller in der Erdrinde vorkommenden Mineralien sind aus kolloidchemischen Prozessen hervorgegangen. Die Ausgangsprodukte der keramischen Industrie bilden Stoffe im kolloiden Verteilungszustand. Typische „Kolloidindustrien" sind weiter die Seifen- und Waschmittel-, Farben-, Gummi- und Textilindustrie.

6.6.7.2.1. *Einteilung der Kolloide.* Man unterscheidet Dispersoid-' Mizell- und Molekülkolloide.

Alle Kolloide, die durch Dispergieren (Verteilen) kompakter fester oder flüssiger Stoffe oder durch Aggregierung der niedermolekularen Bausteine dieser Stoffe in einem flüssigen Dispersionsmittel, in dem diese Stoffe *unlöslich* sind, gewonnen werden können, heißen

Dispersoid- oder *Phasenkolloide*. Die Lyosole sind stets polydispers und meist niederviskos; die Kolloidteilchen in ihnen sind nur aufgrund elektrischer Oberflächenladungen oder adsorbierter grenzflächenaktiver Verbindungen existenzfähig. Zu dieser Gruppe gehören z. B. die meisten anorganischen Hydrosole von Metallen, Sulfiden, Oxiden oder Hydroxiden („eisenhaltiges Wasser") sowie die wäßrigen kolloiden Emulsionen.

Als *Mizell-* oder *Assoziationskolloide* bezeichnet man solche Kolloide, die beim *Lösen* bestimmter niedermolekularer organischer Stoffe in geeigneten Lösungsmitteln oberhalb einer bestimmten Konzentration, der kritischen Konzentration der Mizellbildung, entstehen, wobei sich aber keine echten Lösungen bilden, sondern größere Aggregate aus vielen polaren Einzelmolekülen bzw. -ionen, die durch zwischenmolekulare Wechselwirkungskräfte zusammengehalten werden (↑ 6.6.4.1.2.). Die Lyosole sind meist polydispers und höherviskos; die Mizellen sind oft elektrisch geladen. Beispiele sind die Seifensole und andere höherkonzentrierte „Tensidlösungen".

Kolloide, die aus Einzelmolekülen bestehen, heißen *Molekülkolloide*. Die Sole solcher Kolloide sind stets polymolekular und je nach der Konzentration nieder- bis hochviskos, wobei sie häufig nichtnewtonsches Fließverhalten zeigen. Zu den Molekülkolloiden gehören z. B. alle Eiweißstoffe, auch Gelatine, alle Kohlehydrate wie Dextrin und Stärke und viele synthetische hochpolymere Stoffe.

Hinsichtlich ihrer Gestalt unterscheidet man isometrische und anisometrische Kolloide. Die Teilchen der *isometrischen* oder *Sphärokolloide* haben nach allen drei Raumrichtungen etwa die gleiche Ausdehnung, z. B. Rauch- oder Nebelteilchen, Ruß in schwarzen Farben oder die Stärketeilchen. *Anisometrische Kolloide* bestehen aus Teilchen mit unterschiedlichen räumlichen Ausdehnungen. Häufige Formen sind die *Laminarkolloide* von blättchen-, schicht- oder lamellenförmiger Gestalt, z. B. Glimmer oder die silicatischen Tonmineralien Montmorillonit, Kaolin oder Bentonit sowie bestimmte Mizellkolloide, und die *Linear-* oder *Fibrillarkolloide* mit stäbchen- oder fadenförmiger Gestalt, z. B. viele Molekülkolloide.

6.6.7.2.2. *Lyophobe und lyophile Kolloide*. Infolge der feinen Verteilung der Flüssigkeitströpfchen oder Feststoffpartikel in kolloiden Emulsionen und Suspensionen handelt es sich bei diesen um Systeme mit ausgedehnten inneren Grenzflächen. Sie sind dadurch bedeutend energiereicher als die gleichen stofflichen Systeme im kompakten Zustand. Wenn trotzdem der meist *metastabile kolloiddisperse Zustand* über lange Zeiten aufrechterhalten bleiben kann, müssen stabilisierende Faktoren wirksam sein. Aufhebung dieser Wirkung führt immer zu einer *Koagulation* oder *Flockung*, d. h. zu einer Vereinigung der Kolloidteilchen zu größeren Aggregaten, in denen aber die einzelnen Teilchen im allgemeinen zunächst noch durch dünne Filme aus dem Dispersionsmittel oder durch Adsorptionsschichten voneinander getrennt sind. Aus diesem Zustand ist in manchen Fällen eine Rückführung in die kolloiddisperse Verteilung möglich (*Redispergierung* oder *Peptisation*). Beim Zerreißen der Zwischenschichten erfolgt je nach Art des Systems und den Bedingungen mehr oder weniger schnell *Koaleszenz*, d. h. eine irreversible Vereinigung der Kolloidteilchen zu grobdispersen Aggregaten oder kompakten Phasen unter Abnahme der freien Enthalpie des Systems.

Kolloide, bei denen die Stabilität der entsprechenden Lyosole im wesentlichen durch gleichsinnige elektrische Ladungen oder Tensidadsorptionsschichten bewirkt wird, nennt man *lyophobe* (bei Wasser *hydrophobe*) *Kolloide*. Ursprung der Ladungen ist entweder eine Eigendissoziation der die Kolloidteilchen aufbauenden Moleküle, z. B. $\{Fe(OH)_3\}_n \rightarrow \{(Fe(OH)_3)_{n-x} \cdot (Fe(OH)_2^+)_x\}^{x+} + x\ OH^-$ oder $\{R—COONa\}_n \rightarrow \{(R—COONa)_{n-x} \times (R—COO^-)_x\}^{x-} + x\ Na^+$, oder eine Adsorption von Ionen an den Teilchenoberflächen, z. B. $\{(AgJ)_n \cdot (J^-)_x\}^{x-}$. Derartige Sole zeigen eine ausgeprägte Empfindlichkeit gegenüber Zusätzen entweder von Fremdelektrolyten oder von entgegengesetzt geladenen Kolloiden; in der Praxis wird davon z. B. zur Wasser- oder Abwasserreinigung durch Ausflockung kolloider Schwebstoffe Gebrauch gemacht. Typische lyophobe Kolloide sind z. B. sehr viele anorganische Dispersoidkolloide sowie die dispersen Bestandteile von Emulsionen. Bei letzteren sind für die Stabilität Adsorptionsschichten aus grenzflächenaktiven Stoffen maßgebend (↑ 6.6.7.2.3.).

Es gibt auch eine große Gruppe kolloider Dispersionen in flüssigen Dispersionsmitteln, in denen die dispersen Einzelteilchen in einem *thermodynamisch stabilen Zustand* enthalten sind. Die dispergierten Teilchen werden hier im wesentlichen durch eine *Lyosphäre* (Solvathülle) aus orientiert angelagerten Molekülen des Lösungsmittels stabilisiert, analog etwa zu den Verhältnissen in Elektrolytlösungen (↑ 6.3.4.1.). Typische Beispiele dieser *lyophilen* (*hydrophilen*) *Kolloide* sind z. B. die Kieselsäure oder Oxidhydrate gegenüber Wasser als Dispersionsmittel, viele Eiweißstoffe als hydrophile und viele organische Hochpolymere als organophile Kolloide. Eine wichtige Eigenschaft der Sole lyophiler Kolloide ist ihre ausgeprägte Empfindlichkeit gegenüber anderen Dispersionsmitteln, die durch stärkere Wechselwirkungskräfte auf die Lösungsmittelmole-

küle eine *Desolvatation* der Kolloidteilchen bewirken und diese damit zur Ausflockung bringen, z. B. „Auswaschen" gebrauchter Farbpinsel mit ungeeigneten Lösungsmitteln. Bestimmte lyophile Kolloide, vor allem aus der Gruppe der Molekülkolloide, können auch die Sole lyophober Kolloide stabilisieren. Ihre Wirkung als *Schutzkolloide* beruht darauf, daß sie als stark solvatisierte Teilchen die lyophoben Kolloidteilchen umhüllen und damit eine stabilisierende Wechselwirkung zum Dispersionsmittel vermitteln.

6.6.7.2.3. *Emulsionen, Suspensionen, Schäume.* Die praktisch wichtigsten Kombinationen aus einem flüssigen Dispersionsmittel und flüssigen dispersen Bestandteilen sind die *Emulsionen* aus Wasser und organischen, in Wasser nur sehr begrenzt löslichen Flüssigkeiten, die man in diesem Zusammenhang zusammengefaßt als Öle bezeichnet. In Natur und Technik enthalten beide *Emulsionstypen*, die *Öl-in-Wasser-Emulsionen* (O/W-Emulsionen) und die *Wasser-in-Öl-Emulsionen* (W/O-Emulsionen), stets kolloid- und grobdisperse Anteile gleichzeitig. Welcher der beiden Typen sich beim Zusammenbringen, z. B. durch sehr starkes Rühren oder Schlagen, der beiden Stoffe bildet, hängt von den Mengenverhältnissen und der Temperatur, vor allem aber auch von der Anwesenheit grenzflächenaktiver Stoffe (↑ 6.6.4.1.2.) ab, die man hier als *Emulgatoren* bezeichnet. Grundsätzlich sind Emulsionen nur dann längere Zeit beständig, wenn die dispergierten Tröpfchen durch mindestens monomolekulare Grenzflächenfilme aus Tensidmolekülen stabilisiert sind. Bei O/W-Emulsionen müssen die Emulgatoren in Wasser löslich sein und durch ihren unpolaren, hydrophoben („oleophilen", „lipophilen") Molekülteil chemische Ähnlichkeit mit dem dispergierten Öl besitzen, z. B. Fettsäurederivate oder eiweißartige Stoffe wie Kasein und Albumin als natürliche Emulgatoren in der tierischen Milch oder in der Butter. Weitere solcher Emulsionen sind u. a. Kautschuklatex, Latexfarben, manche Klebstoffe, kosmetische Salben, Tagescrems, medizinische Emulsionen, Pflanzenschutz- und Schädlingsbekämpfungsmittel. Große Bedeutung haben auch wäßrige Emulsionen bestimmter niedermolekularer organischer Verbindungen wie Styrol oder Vinylchlorid, die in dieser Form gut polymerisiert werden können (*Emulsionspolymerisation*). Bei W/O-Emulsionen müssen die Emulgatoren dagegen im Öl löslich sein und mit ihren polaren, hydrophilen Gruppen mit dem dispergierten Wasser über H-Brücken in starke Wechselwirkungen treten können, z. B. Ester und Ether höhermolekularer Fettalkohole, Sterine. Zu solchen Emulsionen gehören u. a. das Roherdöl, Margarine, Mayonnaise, viele Salben und Nachtcrems.

Auch viele Feststoffdispersionen in flüssigen Dispersionsmitteln, d. h. *Suspensionen*, werden durch Tenside nach dem gleichen Wirkungsmechanismus stabilisiert wie Emulsionen. Beispiele sind Anstrichstoffe mit kolloiden Farbpigmenten in Ölen oder Lackgrundstoffen.

Die Bildung und Beständigkeit von *Schäumen* als Verteilungen von Gasen, in erster Linie Luft, in Flüssigkeiten beruhen ebenfalls meist auf der Ausbildung von stabilisierenden Oberflächenfilmen grenzflächenaktiver Stoffe auf den Flüssigkeitslamellen. Gute *Schäumer* bzw. *Schaumstabilisatoren* sind z. B. Seifen, Saponine, Ölsäure oder ätherische Öle. Die Tensidmoleküle sind mit ihren polaren Gruppen in der flüssigen Phase verankert und mit ihren Kohlenwasserstoffmolekülteilen zur Gasphase hin gerichtet. Schäume spielen in der Technik z. B. beim Waschvorgang, bei der Flotation, als Schaumstoffe, Schaumglas und Schaumbeton eine wichtige Rolle. Die Beseitigung bzw. Verhinderung unerwünschter, aber bei vielen technologischen Prozessen auftretender Schäume, etwa bei der Papierherstellung, bei technischen Elektrolysen, beim Transport von Flüssigkeiten mit Pumpen oder in Abwässern, kann mit *Schaumbrechern* erfolgen. Dies sind Tenside, die an der Phasengrenze Flüssigkeit-Gas noch stärker angereichert werden als die Schaumstabilisatoren, selbst aber keinen stabilen Schaum bilden. Für wäßrige Schäume besonders gut schaumbrechende Wirkung haben z. B. höhere Alkohole oder Alkylpolysiloxane (Siliconöle). In flüssiger Form auf die Schäume gegeben, spreiten diese Stoffe auf den Flüssigkeitslamellen unter Verdrängung der ursprünglichen Oberflächenfilme. Dabei reißen sie auch noch einen Teil der Flüssigkeit der Lamellen mit, wodurch diese zum Platzen kommen.

6.6.7.2.4. *Koagulations- und Kondensationsstrukturen, Gele.* Durch Aufhebung der stabilisierenden Faktoren oder durch Verdunstung des Dispersionsmittels gehen alle Arten flüssiger Feststoffdispersionen in einen makroskopisch mehr oder weniger festen, oft pastenförmigen, stets aber leicht plastisch verformbaren Zustand über. Der Zusammenhalt der Kolloidteilchen wird entweder durch dünne Zwischenschichten des Dispersionsmittels oder durch zwischenmolekulare Wechselwirkungskräfte zwischen den Bausteinen der Teilchen bewirkt. Je nach der Geschwindigkeit und sonstigen Bedingungen, unter denen die Koagulation erfolgt, bilden sich *ungeordnete* oder *geordnete Koagulationsstrukturen*, die sich in ihren Eigenschaften unterscheiden. Von grundsätzlicher Bedeutung für die Bildung geordneter Strukturen und z. B. deren mechanische Eigenschaften ist die Form der Kolloidteilchen. Nicht nur anisometrische,

z. B. blättchenförmige Teilchen wie etwa bei V_2O_5- oder nadelförmige Teilchen wie bei γ-Fe_2O_3-Partikeln, sondern auch kugelförmige Kolloidteilchen wie feinste SiO_2-Partikel können zu einer „Erstarrung" des Lyosols unter Ausbildung dreidimensional vernetzter Raumstrukturen führen. Bei der geordneten Koagulation von kugelförmigen Teilchen bilden sich Kettenstrukturen. In vielen Fällen beobachtet man dann bei mechanischer Druckbeanspruchung eine Verfestigung, indem noch vorhandenes Dispersionsmittel von den Teilchen abgeschert und aus den Zwischenräumen herausgepreßt wird. Bei Aufhören der Belastung wird diese Flüssigkeit wieder angelagert, wodurch das System wieder weicher wird und häufig pastenförmig zerfließt (Stärkebrei, Teige oder Fensterkitte; Treten von nassem Sand). Diese reversible Zustandsänderung bezeichnet man als *Dilatanz*.

Die geordneten Koagulationsstrukturen anisometrischer Kolloide zeigen die ebenfalls reversible Erscheinung der *Thixotropie*. Bei mechanischer Bewegung werden diese Systeme leichtflüssiger, weil der relativ schwache Zusammenhalt zwischen den Teilchen durch die Krafteinwirkung mehr oder weniger aufgehoben wird. Die Thixotropie ist von erheblicher praktischer Bedeutung, z. B. für die Verarbeitung von Ton- oder anderen Silicatmaterialien in der keramischen Industrie, von Ölfarben und Klebstoffen oder für das Verhalten von Druckfarben in Druckmaschinen.

Beim Zerreißen der Flüssigkeitsschichten zwischen den Kolloidteilchen in Koagulationsstrukturen oder durch Ablauf chemischer Reaktionen in ihnen (Erhärten von Baustoffen) können solche Strukturen unter Koaleszenz in *Kondensationsstrukturen* übergehen, in denen die Teilchen, manchmal unter Ausbildung ausgedehnter Hohlräume oder Poren, zu *Xerogelen* zusammenwachsen. Diese sind manchmal von horn- oder kunststoffartiger Natur, in anderen Fällen von kristallinem Aufbau, z. B. das Silicagel. Ihre mechanische Festigkeit ist stets wesentlich größer als die von Koagulationsstrukturen.

Auch viele lyophile Molekülkolloide bilden bei Flüssigkeitsentzug vernetzte Raumstrukturen aus, durch die sie als *Lyogele* elastisch verformbare, weiche Massen bilden. Sie können oftmals noch beträchtliche Mengen an Dispersionsmittel enthalten, z. B. etwa 96% Wasser im Gelatineguß von Obsttorten oder im Glaskörper des Auges. Daher verhalten sich Lyogele ähnlich wie die betreffenden Lyosole, z. B. hinsichtlich der Wanderungsfähigkeit niedermolekularer Verbindungen oder Ionen in ihnen. Durch Deformation und weiteren Flüssigkeitsentzug kann man zu Fasern, Filmen, Folien oder anderen Formkörpern gelangen.

Charakteristisch für viele hochmolekulare Natur- und Kunststoffe ist ihr *Quellungsvermögen*; unter Aufnahme von geeignetem Dispersionsmittel können diese Xerogele in Lyogele, bei *unbegrenzter Quellbarkeit* bis in den Zustand der Lyosole übergehen (Gummilösung, Gelatinelösung, PVC-Kleber). In vielen Fällen ist die Umwandlung Xerogel-Lyogel-Lyosol reversibel; irreversible Xerogelbildung tritt z. B. beim Hühnereiweiß beim Kochen und allgemein bei Ablauf chemischer Reaktionen mit den Makromolekülen auf. Viele Naturprodukte haben eine *begrenzte Quellbarkeit*, z. B. Samenkörner oder viele Nahrungsmittel wie Teig- und Backwaren.

7. Stoffwandlung

7.1. Grundbegriffe

7.1.1. Chemische Gleichgewichte

7.1.1.1. Stoffwandlung durch chemische Reaktion. Jeder Stoff ist unter den jeweils gegebenen Bedingungen (Druck, Temperatur) durch seine Zusammensetzung, seine Struktur und seine physikalischen und chemischen Eigenschaften gekennzeichnet. Von Stoffwandlungen im weitesten Sinne kann man sprechen, wenn z. B. durch Änderung des Aggregatzustands (↑ 6.1.1.1.) oder stärkere Wechselwirkungen eines Stoffs die Natur bzw. Stärke der Bindungen und die Anordnung der Atome bzw. Moleküle verändert werden. Als Stoffwandlungen im chemischen Sinne, d. h. als *chemische Reaktionen*, werden im allgemeinen nur solche Vorgänge angesehen, bei denen durch Elektronenumgruppierung unter Struktur-, Eigenschafts-, meist auch Zusammensetzungsänderung neue Stoffe entstehen.

7.1.1.2. Stöchiometrie von chemischen Reaktionen. Zur Beschreibung einer chemischen Reaktion benötigt man Angaben über die miteinander reagierenden Ausgangsstoffe, auch als *Edukte* oder *Reaktanten* bezeichnet, und die entstehenden Produkte, über die Reaktionsbedingungen und die die Reaktion begleitenden Energieumsätze.

Die an der Reaktion beteiligten Stoffe werden in der *stöchiometrischen* oder *Bruttoreaktionsgleichung* mit ihren Formeln angegeben und in ihrem zahlenmäßigen Verhältnis zueinander durch die *stöchiometrischen Faktoren* $|\nu_i|$ charakterisiert, z. B. $2\,HCl + Zn \rightarrow ZnCl_2 + H_2$ oder allgemein $|\nu_A|\,A + |\nu_B|\,B \rightarrow |\nu_C|\,C + |\nu_D|\,D$. Bei der Aufstellung einer Reaktionsgleichung sind folgende Prinzipien einzuhalten:

Alle umgesetzten und entstehenden Stoffe werden berücksichtigt.

Stoffbilanz: Die Summe der Atome auf der Ausgangsseite ist gleich der Summe der Atome auf der Produktseite.

Ladungsbilanz: Die Summe der Ladungen muß für beide Seiten der Gleichung identisch sein.

Die stöchiometrischen Faktoren sollen als die kleinstmöglichen ganzen Zahlen auftreten.

Die stöchiometrischen Faktoren $|\nu_i|$ sind die Absolutwerte der *Stöchiometriezahlen* ν_i, die vorzeichenbehaftet sind. Für Ausgangsstoffe sind sie negativ, für Produkte positiv. Für die Reaktion des Zinks mit Salzsäure ist $\nu_{HCl} = -2$, $\nu_{Zn} = -1$, $\nu_{ZnCl_2} = 1$ und $\nu_{H_2} = 1$.

Eine so aufgestellte Reaktionsgleichung läßt außer der qualitativen Information, welche Stoffe zur Bildung eines bestimmten Produkts nötig sind, weitgehende quantitative Aussagen zu, die Gegenstand der *Stöchiometrie* sind [*stoicheion* griech., Elementarbestandteil; *metrein* griech., messen]. Dabei wird ausgenutzt, daß die chemischen Formeln in der Reaktionsgleichung eine Doppelbedeutung haben. Unter der Formel HCl kann entweder ein konkretes Molekül HCl als elementares Teilchen oder die Objektmenge 1 mol HCl, d. h. die Gesamtheit von $6{,}0225 \cdot 10^{23}$ Molekülen HCl, verstanden werden. Das bedeutet aber auch, daß entsprechend der Reaktionsgleichung $2\,HCl + Zn \rightarrow ZnCl_2 + H_2$ zwei Moleküle bzw. 2 mol HCl auf ein Atom bzw. 1 mol Zn einwirken müssen, um eine Formeleinheit bzw. 1 mol $ZnCl_2$ zu erhalten. Werden 3 mol HCl mit der ausreichenden Menge Zn umgesetzt, können dementsprechend maximal 1,5 mol $ZnCl_2$ gebildet werden.

Die stöchiometrischen Faktoren $|\nu_i|$ in der allgemeinen Reaktionsgleichung geben also an, in welchen relativen Mengen die Ausgangsstoffe miteinander reagieren und welche relativen Mengen an Produkten bei vollständigem Umsatz zu erwarten sind.

Das heißt, daß zur Umsetzung einer bestimmten Menge n_A des Stoffs A eine Menge $n_B = n_A|\nu_B|/|\nu_A|$ des Stoffs B benötigt wird, um die Menge $n_C = n_A|\nu_C|/|\nu_A|$ des Stoffs C zu bilden. Die Mengenangabe erfolgt dabei üblicherweise als Objektmenge n_i mit der Einheit mol.

Die entsprechenden umgesetzten *Massen* m_i können aus n_i und der *Molmasse* M_i nach $m_i = n_i M_i$ erhalten werden. Im obigen Beispiel ergeben 109 g HCl (3 mol) maximal 204 g $ZnCl_2$ (1,5 mol).

Sind Gase an der Reaktion beteiligt, läßt sich deren Volumen sehr leicht unter der Annahme berechnen, daß sie sich wie ein ideales Gas verhalten. Das *Molvolumen* V_0 eines idealen Gases bei 0 °C und dem Druck von $101{,}325 \cdot 10^3$ Pa (1 atm) beträgt $22{,}4\,l\,mol^{-1}$. Bei der Umsetzung von 109 g HCl mit Zink werden 3 g H_2 frei, die bei 0 °C ein Volumen von 33,6 l einnehmen. Die Umrechnung des Volumens eines idealen Gases i auf andere Temperaturen und Drücke erfolgt mit der Zustandsgleichung idealer Gase: $pv_i = n_i RT$ (↑ 6.2.2.1.5.).

Die meisten Reaktionen laufen nicht zwischen reinen Stoffen ab, sondern gehen aus von bzw. führen zu *Mischungen* oder *Lösungen*. In diesen Fällen muß die Änderung der *Zusammensetzung der Mischphasen* infolge der Reaktion quantitativ angegeben werden. Die dazu erforderlichen Zusammensetzungsvariablen Molenbruch, Massenbruch, Molalität bzw. Konzentration und ihre Umrechnung ineinander sind in 6.1. angegeben; für Gasmischungen kann der *Partialdruck* $p_i = x_i p$, d. h. das Produkt aus Molenbruch der Gaskomponente i und Gesamtdruck p, anstelle des Molenbruchs x_i verwendet werden.

7.1.1.3. Chemisches Gleichgewicht und Massenwirkungsgesetz. Während bei der Reaktion $2\,HCl + Zn \rightarrow ZnCl_2 + H_2$ praktisch vollständige Umsetzung stattfindet (↑ 7.1.1.5.), liegt z. B. bei der Bildung von Ester E und Wasser aus Säure S und Alkohol A im Gleichgewicht, unabhängig von den gewählten Ausgangskonzentrationen, ein Gemisch aller Reaktionsteilnehmer vor. Das Verhältnis ihrer Aktivitäten ist dann konstant; es läßt sich nach dem Massenwirkungsgesetz (MWG) (↑ 6.1.2.4.3.) zu $K = a_E a_{H_2O}/(a_S a_A)$ angeben.

Für die allgemeine Reaktionsgleichung $|\nu_A|\, A + |\nu_B|\, B \rightleftharpoons |\nu_C|\, C + |\nu_D|\, D$ führt das MWG zu $K = a_C^{|\nu_C|} a_D^{|\nu_D|}/(a_A^{|\nu_A|} a_B^{|\nu_B|})$. Die *thermodynamische Gleichgewichtskonstante K* hat für jedes p und T einen für jede Reaktion spezifischen Wert.

Bei der Aufstellung des MWG wird von der stöchiometrischen Reaktionsgleichung ausgegangen; den Nenner des Bruchs bilden immer die Aktivitäten der Ausgangsstoffe (↑ 6.1.2.). Das MWG drückt aus, daß eine Reaktion solange abläuft, bis das Verhältnis der Aktivitäten aller Reaktionsteilnehmer seinen durch das MWG bestimmten Gleichgewichtswert angenommen hat. Die Richtung der Gleichgewichtseinstellung hängt dabei von den Ausgangsaktivitäten a_i^0 aller Reaktionsteilnehmer ab. Ist der Quotient $Q^0 = (a_C^{|\nu_C|})^0\, (a_D^{|\nu_D|})^0/[(a_A^{|\nu_A|})^0\, (a_B^{|\nu_B|})^0]$ größer als K, dann wird das Gleichgewicht durch Verbrauch der Stoffe C und D erreicht. Bei $Q^0 < K$ werden zur Gleichgewichtseinstellung die Stoffe C und D aus A und B gebildet; für $Q^0 = K$ ist das System gerade im Gleichgewicht; es findet kein Stoffumsatz statt. Die im MWG auftretenden *Aktivitäten* der einzelnen Stoffe sind einer direkten Messung nicht zugänglich. Deshalb werden in der Praxis häufig nur die *konventionellen Gleichgewichtskonstanten* K_c, K_p oder K_x bestimmt und tabelliert. Sie sind analog definiert, enthalten aber statt der Aktivitäten die Zusammensetzungsvariablen Konzentration c_i, Partialdruck p_i oder Molenbruch x_i, z. B. $K_x = x_C^{|\nu_C|} x_D^{|\nu_D|}/(x_A^{|\nu_A|} x_B^{|\nu_B|})$. Da die Zusammensetzungsvariablen analytisch bestimmt sind und nicht die durch die zwischenmolekularen Wechselwirkungen beeinflußten Aktivitäten wiedergeben, sind die konventionellen Gleichgewichtskonstanten im Unterschied zu den thermodynamischen für gegebenes p und T nicht oder nicht streng konstant, sondern abhängig von der Zusammensetzung der Reaktionsmischung. Der Zusammenhang zwischen den Aktivitäten und den Zusammensetzungsvariablen wird über die *Aktivitätskoeffizienten f* hergestellt (↑ 6.1.2.4.1.).

7.1.1.4. Beispiele für chemische Gleichgewichte. Bei Kenntnis der Reaktionslaufzahl λ (↑ 6.1.2.4.3.) läßt sich die Lage chemischer

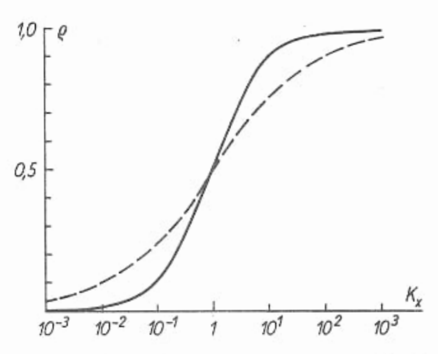


Abb. 7.1.1-1 Zusammenhang zwischen Gleichgewichtskonstante K_x und Reaktionsgrad ϱ für die Gleichgewichte; $A \rightleftharpoons B$ (*durchgezogene Linie*) und $A + B \rightleftharpoons C + D$ (*gestrichelte Linie*)

Gleichgewichte besonders anschaulich durch den *Reaktionsgrad* ϱ beschreiben. Er ergibt sich als Quotient aus dem Umsatz λ_{equ} bis zum Gleichgewicht und dem maximalen Umsatz λ_{max}, der bei vollständigem Verbrauch des im Unterschuß vorliegenden Ausgangsstoffs A erreicht ist: $\varrho = \lambda_{equ}/\lambda_{max}$ mit $\lambda_{max} = n_A^0/|\nu_A|$.

Der besonders einfache Fall der Umwandlung eines Stoffs in einen anderen, $A \rightleftharpoons B$, liegt z. B. bei der Keto-Enol-Tautomerie, $R{-}CO{-}CH_2{-}R' \rightleftharpoons R{-}C(OH){=}CH{-}R'$, vor. Der Reaktionsgrad wird hier gemäß $\varrho = K_x/(1 + K_x)$ nur vom Wert K_x der Gleichgewichtskonstante bestimmt und zeigt seine stärkste Änderung im Bereich $K_x = 10^{-1}$ bis 10^1 (Abb. 7.1.1-1).

Der Zerfall eines Stoffs A in zwei Produkte gemäß $A \rightleftharpoons B + C$ ist besonders häufig als Dissoziationsgleichgewicht von Säuren und Salzen in Lösung realisiert. Der Reaktionsgrad wird dann als Dissoziationsgrad α bezeichnet. Aus $K_c = \alpha^2 c_A^0/(1 - \alpha)$ ist ersichtlich, daß der Dissoziationsgrad mit steigendem K_c und fallender Ausgangskonzentration c_A^0 wächst; für $K_c/c_A^0 = 10$ ergibt sich z. B. $\alpha = 0{,}92$. Selbst sehr schwach dissoziierende Stoffe

können deshalb bei ausreichend niedriger Konzentration vollständig dissoziieren. So liegt das Salz $[(C_6H_5)_3C^+]$ $[SbCl_6^-]$ trotz seiner ziemlich kleinen Dissoziationskonstante von $1{,}4 \cdot 10^{-4}$ mol l^{-1} in 10^{-5} molarer Lösung in CH_2Cl_2 weitgehend in freien Ionen vor.

Doppelte Umsetzungen in homogener Phase gemäß $A + B \rightleftharpoons C + D$ liegen bei Substitutionsreaktionen vor, wie etwa der Veresterung einer Säure (↑ 7.1.1.3.), $CH_3COOH + C_2H_5OH \rightleftharpoons CH_3COOC_2H_5 + H_2O$. Für einen äquivalenten Einsatz von A und B, d. h. für $x_A^0 = x_B^0$, ergibt sich $\varrho = \sqrt{K_x}/(1 + \sqrt{K_x})$ im Gleichgewicht (↑ Abb. 7.1.1-1). Bei beliebigen Ausgangsmolenbrüchen x_A^0, x_B^0 hängt der Reaktionsgrad sowohl von derem Verhältnis als auch von K_x ab: $K_x = \alpha^2/[(x_A^0/x_B^0) - \varrho]\,(1 - \varrho)$. Jede Erhöhung der Objektmenge eines Ausgangsstoffs gegenüber der des anderen erhöht den Reaktionsgrad, ein Zuwachs der Anfangsobjektmenge eines Produkts erniedrigt ihn (Abb. 7.1.1-2).

Für Reaktionen in der Gasphase wird das MWG mit den Partialdrücken p_i formuliert. Bei der Berechnung der Partialdrücke aus der Gaszusammensetzung nach $p_i = x_i p = n_i p/\sum n_j$ müssen in der Summe von n_j auch alle an der Reaktion nicht teilnehmenden Fremdgase berücksichtigt werden. So ergibt sich für das Dissoziationsgleichgewicht $N_2O_4 \rightleftharpoons 2\,NO_2$ mit $K_p = 7{,}7 \cdot 10^4$ Pa (0,76 atm) bei 50 °C für $p = 1{,}01 \cdot 10^5$ Pa (1 atm) und einem Einsatz von 1 mol reinem N_2O_4 ein Dissoziationsgrad $\alpha = 0{,}4$. Bei „Verdünnung" der Stickstoffoxide durch einen zehnfachen Überschuß eines Inertgases wächst α auf 0,749.

Bei heterogenen Gleichgewichten komplizieren

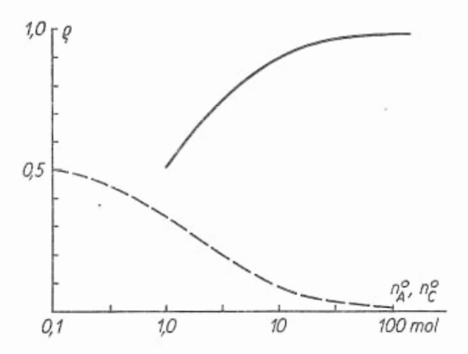


Abb. 7.1.1-2 Zusammenhang zwischen der Ausgangsobjektmenge von Reaktionsteilnehmern am Gleichgewicht $A + B \rightleftharpoons C + D$ und dem Reaktionsgrad bei $K_x = 1$; *durchgezogene Kurve* für $n_A^\ominus$ bei $n_B^\ominus = 1$ mol, $n_C^\ominus = n_D^\ominus = 0$; *gestrichelte Kurve* für $n_C^\ominus$ bei $n_A^\ominus = n_B^\ominus = 1$ mol, $n_D^\ominus = 0$

sich meist die Verhältnisse, da nicht mehr für alle Reaktionspartner derselbe Standardzustand zugrundegelegt werden kann. Treten z. B. Gase als Produkte von Reaktionen in kondensierter Phase auf, dann gehen sie meist mit ihrem Partialdruck p_i im Gasraum in das MWG ein, das dann mit unterschiedlichen Zusammensetzungsvariablen für die einzelnen Reaktionsteilnehmer formuliert werden muß.

Sind reine Stoffe an heterogenen Gleichgewichten beteiligt, ergibt sich dadurch eine Vereinfachung, daß deren Aktivität $a = 1$ zu setzen ist. Somit wird das Löslichkeitsgleichgewicht schwerlöslicher Salze, z. B. gemäß $AgCl_{(fest)} \rightleftharpoons Ag^+ + Cl^-$, wegen $a_{AgCl} = 1$ schon durch das Löslichkeitsprodukt $L_{AgCl}^{(20\,°C)} = c_{Ag^+} \cdot c_{Cl^-} = 1{,}6 \cdot 10^{-10}$ mol^{2}l^{-2} hinreichend beschrieben. Jede Erhöhung der Konzentration des einen Ions durch gleichionigen Zusatz, z. B. von Natriumchlorid, führt solange zur Ausscheidung von Silberchlorid, bis der Gleichgewichtswert L_{AgCl} wieder erreicht ist.

Für die heterogene Reaktion $2\,HCl + Zn \rightleftharpoons ZnCl_2 + H_2$, die als Reaktion in wäßriger Lösung besser in der Form $2\,H^+ + Zn \rightleftharpoons Zn^{2+} + H_2$ formuliert wird, ergibt sich mit $a_{Zn} = 1$ die gemischte Gleichgewichtskonstante $K_{c,p} = c_{Zn^{2+}}p_{H_2}/c_{H^+}$. Bei Durchführung der Reaktion im offenen Gefäß bedeutet das aber, daß beim Entweichen des Wasserstoffs in die Atmosphäre mit ihrem ungeheuren Überschuß an Fremdgas der Molenbruch x_{H_2} und damit der Partialdruck p_{H_2} immer praktisch Null bleiben. Die Umsetzung verläuft deshalb bis zum vollen Verbrauch der Ausgangsstoffe, ohne daß das Gleichgewicht erreicht wird.

Von gekoppelten Gleichgewichten spricht man, wenn das Endprodukt einer Reaktion zugleich Ausgangsstoff einer Folgereaktion ist, z. B. bei der Dissoziation einer zweiwertigen Säure in Wasser bei 25 °C: $H_2S \rightleftharpoons H^+ + HS^-$ mit $K_c' = 1{,}1 \cdot 10^{-7}$ mol l^{-1} und $HS^- \rightleftharpoons H^+ + S^{2-}$ mit $K_c'' = 1{,}0 \cdot 10^{-14}$ mol l^{-1}. Für den Bruttovorgang $H_2S \rightleftharpoons 2\,H^+ + S^{2-}$ ist die Gleichgewichtskonstante als Produkt der Einzelkonstanten: $K_c = K_c'K_c'' = 1{,}1 \cdot 10^{-21}$ mol^{2}l^{-2}.

7.1.1.5. Beeinflussung chemischer Gleichgewichte. Qualitative Voraussagen zur Richtung der Gleichgewichtsverschiebung gestattet das „Prinzip des kleinsten Zwangs" nach LE CHATELIER und BRAUN:

Übt man auf ein System im thermodynamischen Gleichgewicht durch Änderung von Druck, Temperatur oder Zusammensetzung einen Zwang aus, so läuft die Reaktion in der Richtung ab, die den Zwang abbaut.

7.1.1.5.1. *Änderung der Zusammensetzung.* Für die allgemeine Reaktion $|\nu_A|\,A + |\nu_B|\,B \rightleftharpoons |\nu_C|\,C + |\nu_D|\,D$ muß im Gleichgewicht der Ausdruck $a_C^{|\nu_C|}a_D^{|\nu_D|}/(a_A^{|\nu_A|}a_B^{|\nu_B|})$ als Ganzes kon-

stant und gleich K sein. Daraus folgt, wie in 7.1.1.4. schon an Beispielen belegt, daß sich die einzelnen Aktivitäten bei Konstanz von K relativ zueinander ändern können. Die Umsetzung von Zink mit Salzsäure zeigte, daß durch laufende Entfernung eines flüchtigen Produkts aus dem Gleichgewicht die Reaktion im Grenzfall bis zum vollständigen Verbrauch der Ausgangsstoffe gehen kann, ohne daß das Gleichgewicht erreicht wird. Ähnliche Situationen ergeben sich, wenn eins der Produkte als schwerlöslicher Stoff ausgefällt und abgetrennt wird oder man eins der Produkte chemisch bindet, wie z. B. Wasser bei der Esterbildung durch Zugabe konzentrierter Schwefelsäure:

$NaCl + H_3PO_4 \rightarrow HCl\uparrow + NaH_2PO_4$;

$BaO_2 + H_2SO_4 \rightarrow BaSO_4\downarrow + H_2O_2$;

$CH_3COOH + C_2H_5OH \xrightarrow[H_2SO_4]{} H_3CCOOC_2H_5 + H_2O$.

7.1.1.5.2. *Temperaturänderung.* Die Abhängigkeit der Gleichgewichtskonstante K von der Temperatur bei konstantem Druck ist durch die van't-Hoffsche Reaktionsisobare gegeben (↑ 6.1.2.4.3.). Ihre Integration liefert für eine konstante Standardreaktionsenthalpie bei der absoluten Temperatur T: $\lg K(T) = \lg K(T_0) - (\Delta_R H^\ominus/2{,}303R)\,(1/T - 1/T_0)$; zur Vereinfachung wurde angenommen, daß $\Delta_R H^\ominus$ temperaturunabhängig ist (1. Ulichsche Näherung). Mit steigender Temperatur wächst die Konstante K, wenn die molare Standardreaktionsenthalpie $\Delta_R H^\ominus$ positiv ist, d. h. die Reaktion endotherm verläuft, sie fällt dagegen bei exothermen Reaktionen. Die Reaktion weicht also bei Temperaturerhöhung zur Seite des Wärmeverbrauchs aus. Wie beträchtlich der Temperatureinfluß auf die Gleichgewichtslage sein kann, zeigt die Ammoniaksynthese $N_2 + 3\,H_2 \rightleftharpoons 2\,NH_3$ mit $\Delta_R H^\ominus = -92{,}5$ kJmol^{-1}. Bei 298 K hat die Gleichgewichtskonstante einen Wert von $K_p = 6{,}6 \cdot 10^{-5}$ Pa^{-2}, bei 498 K ist sie auf $K_p = 0{,}2 \cdot 10^{-10}$ Pa^{-2} abgefallen.

7.1.1.5.3. *Druckänderung.* Die Abhängigkeit der Konstanten K_x vom Druck bei konstanter Temperatur folgt dem Gesetz $(\partial \ln K_x/\partial p)_T = -\Delta_R V^\ominus/RT$.

$\Delta_R V^\ominus$, das molare *Standardreaktionsvolumen*, ist die Volumenänderung bei einem Formelumsatz unter Standardbedingungen (↑ 6.1.2.4.3.). Dehnt sich das System bei der Reaktion aus, erhält $\Delta_R V^\ominus$ einen positiven Wert.

Bei *Reaktionen in kondensierten Phasen*, also unter Ausschluß gasförmiger Stoffe, ist das molare Reaktionsvolumen vernachlässigbar klein. Nur bei höchsten Drücken wird eine deutliche Beeinflussung von K_x beobachtet. Für das Gleichgewicht C (Graphit) $\rightleftharpoons$ C (Diamant) z. B. ist $\Delta_R V^\ominus = -1{,}91$ cm^{3}mol^{-1}. Bei dem Normaldruck von $p \approx 10^5$ Pa (1 atm) ergibt sich $K_x = 0{,}43$, d. h., die Bildung von Graphit aus Diamant ist begünstigt. Erst für Drücke $p > 8{,}3 \cdot 10^9$ Pa (> 82200 atm) wird $K > 1$ und damit eine *Diamantsynthese* möglich.

Weit deutlicher ist der Druckeinfluß auf solche Gleichgewichte, an denen Gase beteiligt sind. Bei Druckerhöhung verschiebt sich die Lage des Gleichgewichts zugunsten *der* Seite mit dem geringeren Anteil gasförmiger Stoffe.

Das molare Reaktionsvolumen steht unmittelbar mit den Stöchiometriezahlen der an der Reaktion beteiligten Gase j im Zusammenhang. Nimmt man ideales Verhalten der Gase an, so gilt: $\Delta_R V^\ominus = \sum \nu_j V^\ominus = \sum \nu_j RT/p$. Die Integration der obigen Differentialgleichung ergibt: $\lg K_x(p) = \lg K_x(p_0) - \sum \nu_j \lg (p/p_0)/2{,}303$.

Demzufolge wird für die Ammoniaksynthese wegen $\sum \nu_j = -2$ die Gleichgewichtskonstante K_x durch eine Druckerhöhung um etwa $2 \cdot 10^7$ Pa (200 atm) um den Faktor 10^2 erhöht.

7.1.2. Thermodynamische Beschreibung chemischer Reaktionen

7.1.2.1. Molare freie Standardreaktionsenthalpie und Gleichgewichtskonstante. In 6.1.2.3.2. wurden als Bedingung für den freiwilligen Ablauf von Prozessen negative Werte für die Änderung der freien Energie bzw. freien Enthalpie genannt. Liegen alle Reaktionsteilnehmer X_i mit der Aktivität $a_i^\ominus = 1$ vor, so gilt für den Zusammenhang zwischen der dann zu betrachtenden molaren freien Standardreaktionsenthalpie $\Delta_R G^\ominus$ und der Gleichgewichtskonstante K die Beziehung $\Delta_R G^\ominus = -RT \ln K$ (Abb. 7.1.2-1). Dabei sind drei Fälle zu unterscheiden: $K > 1$ bzw. $\Delta_R G^\ominus < 0$: die Reaktion läuft freiwillig nach rechts; $K = 1$ bzw. $\Delta_R G^\ominus = 0$: Gleichgewicht; $K < 1$ bzw. $\Delta_R G^\ominus > 0$: die Reaktion läuft freiwillig nach links.

Sind die K-Werte nicht bekannt, lassen sich die $\Delta_R G^\ominus$-Werte aus den molaren freien Standardbildungsenthalpien $\Delta_B G^\ominus$ auf analoge Weise

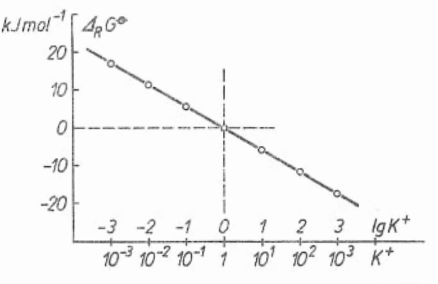


Abb. 7.1.2-1 Zusammenhang zwischen Gleichgewichtskonstante K und molarer freier Standardreaktionsenthalpie $\Delta_R G^\ominus$

berechnen, wie es in 6.1.2.2. für die Ermittlung der molaren Standardreaktionsenthalpien $\Delta_R H^\ominus$ aus den molaren Standardbildungsenthalpien $\Delta_B H^\ominus$ angegeben wurde. Erforderlichenfalls werden die $\Delta_B G^\ominus$-Werte mittels einer geeignet modifizierten Gibbs-Helmholtz-Gleichung berechnet (↑ 6.1.2.3.2.):
$\Delta_B G^\ominus = \Delta_B H^\ominus - T\Delta_B S^\ominus$.

Tab. 7.1.2-2 Molare Standardgrößen einiger Stoffe bei 298,15 K und 101,325 · 10³ Pa (1 atm)

Stoff	$\Delta_B H^\ominus$ kJmol⁻¹	$S^\ominus$ Jmol⁻¹K⁻¹	$\Delta_B G^\ominus$ kJmol⁻¹
H_2	0	130,5	0
O_2	0	205	0
H_2O, gasf.	−242	189	−229
H_2O, flüss.	−286	70	−237
Cl_2	0	223	0
HCl	−92	187	−95
S, rhombisch	0	32	0
SO_2	−297	248	−300
SO_3	−395	256	−370
N_2	0	191,5	0
NH_3	−46	192,5	−17
NO_2	+34	240,5	+52
N_2O_4	+10	304	+98
C, Graphit	0	5,5	0
CO	−110,5	198	−137
CO_2	−393	213,5	−393,5
CH_4	−75	186	−51
C_6H_6	+49	173	+124,5
n-Butan	−125	310	−16
n-Pentan	−146,5	348,5	−8
2-Methylbutan	−154,5	343	−14,5
n-Hexan	−167	387	+0,2
n-Heptan	−188	425	+8,5
Ethylen	+52	219,5	+68
HCN, flüss.	+105	113	+121

Die benötigten Werte sind im allgemeinen tabelliert (Tab. 7.1.2-2); so können für die Bildung von Wasserdampf aus den Elementen bei 298 K und 101,325 · 10³ Pa die folgenden Werte entnommen bzw. berechnet werden:

$$\Delta_B H^\ominus_{H_2O} = \Delta_B H^\ominus_{H_2O} - \Delta_B H^\ominus_{H_2} - 1/2\Delta_B H^\ominus_{O_2} = (-242 - 0 - 0)\ \text{kJmol}^{-1} = -242\ \text{kJmol}^{-1};$$

$$\Delta_B S^\ominus_{H_2O} = S^\ominus_{H_2O} - S^\ominus_{H_2} - 1/2 S^\ominus_{O_2} = (189 - 130{,}5 - 102{,}5)\ \text{Jmol}^{-1}\text{K}^{-1} = -44\ \text{Jmol}^{-1}\text{K}^{-1};$$

$$\Delta_B G^\ominus_{H_2O} = -242\ \text{kJmol}^{-1} - 298\ \text{K} \times (-44)\ \text{Jmol}^{-1}\text{K}^{-1} = (-242 + 13)\ \text{kJmol}^{-1} = -229\ \text{kJmol}^{-1}.$$

Durch das Experiment zugänglich sind die Bildungsenthalpien aus den Verbrennungswärmen (↑ 6.1.2.2.4.), die freien Reaktions- bzw. Bildungsenthalpien aus der Gleichgewichtszellspannung U_{equ} entsprechend konstruierter galvanischer Zellen nach $\Delta_R G = z_r F U_{equ}$ (↑ 7.3.4.), die Standardentropien $S^\ominus$ aus Messungen der Wärmekapazität sowie der Phasenumwandlungsenergien der jeweiligen Stoffe.

Wenn sich die Reaktionsteilnehmer nicht, wie bisher angenommen, im Standardzustand befinden, so ist in Übereinstimmung mit den Ausführungen zum chemischen Potential (s. a. 6.1.2.4.1.) das Standardglied $\Delta_R G^\ominus$ um ein Überführungsglied zur freien molaren Reaktionsenthalpie $\Delta_R G$ zu erweitern:

$$(\Delta_R G)_{p,T} = (\Delta_R G^\ominus)_{p,T} + RT \ln (a_C^{|\nu_C|} a_D^{|\nu_D|} / a_A^{|\nu_A|} a_B^{|\nu_B|}).$$

Eine entsprechende Beziehung gilt für das Arbeiten bei konstantem Volumen und konstanter Temperatur, also für die freie molare Reaktionsenergie $\Delta_R F$. Beide Ausdrücke werden als *van't-Hoffsche Reaktionsisotherme* bezeichnet. Sie beschreibt den Einfluß beliebiger Aktivitäten (Drücke, Konzentrationen) der Reaktionspartner auf die molare freie Reaktionsenthalpie bzw. -energie und vermag somit die Frage nach dem freiwilligen Ablauf einer interessierenden Reaktion unter den jeweils vorgegebenen Bedingungen zu beantworten.

7.1.2.2. Näherungsweise Berechnung thermodynamischer Größen

7.1.2.2.1. *Abschätzung von $\Delta_R H^\ominus$- und $\Delta_R G^\ominus$-Werten.* Die Reaktionsenthalpie summiert über die energetischen Veränderungen beim Ablauf einer Reaktion. Diese lassen sich in günstigen Fällen beschreiben durch Aufspaltung und

Tab. 7.1.2-3 Richtwerte für die mittleren molaren Bindungsenergien E_{X-Y} einiger Bindungen in kJ pro Mol Bindungen

X—Y	E_{X-Y}	X—Y	E_{X-Y}
H—H	435	H—I	300
C—C	345	C—N	300
N—N	160	C—O	350
O—O	145	C—F	460
F—F	160	C—Cl	330
Cl—Cl	245	C—Br	280
Br—Br	190	C—I	230
H—C	415	C=C	615
H—N	390	C≡C	815
H—O	465	C=N	615
H—F	565	C≡N	800
H—Cl	430	C=O	720
H—Br	365		

Neubildung von Bindungen beim Übergang von den Ausgangsstoffen zu den Endprodukten. Bei Kenntnis der mittleren molaren Bindungsenergien E_{X-Y} (Tab. 7.1.2-3) ergibt sich dann $\Delta_R H^\ominus$ näherungsweise zu $\Delta_R H^\ominus \approx -(\sum E_{X-Y(\text{gebildet})} - \sum E_{X-Y(\text{gebrochen})})$. Im Zuge der Reaktion $H_2 + Cl_2 \rightarrow 2\,HCl$ werden eine Bindung H—H (435 $kJmol^{-1}$) und eine Bindung Cl—Cl (245 $kJmol^{-1}$) gebrochen sowie zwei Bindungen H—Cl ($2 \cdot 430\ kJmol^{-1}$) gebildet. Daraus folgt $\Delta_R H^\ominus = -180\ kJmol^{-1}$, d. h., es besteht gute Übereinstimmung mit der molaren Standardbildungsenthalpie $\Delta_B H^\ominus$ von HCl (↑ Tab. 7.1.2-2). Allgemein sind exotherme Reaktionen immer dann zu erwarten, wenn schwache Bindungen gebrochen und starke neugebildet werden.

Für gasförmige organische Verbindungen wurden Systeme entwickelt, die es gestatten, $\Delta_B H^\ominus$-Werte aus Inkrementen I_H für einzelne Strukturelemente näherungsweise zu berechnen (Tab. 7.1.2-4). Für 2-Methylbutan ergäbe sich z. B. $\Delta_R H^\ominus \approx 3I_H(\text{—CH}_3) + I_H(\text{—CH}_2\text{—}) + I_H\left(\rangle\text{CH—}\right) \approx -152\ kJmol^{-1}$. Mit Hilfe so abgeschätzter oder schon tabellierter $\Delta_B H^\ominus$-Werte können dann die molaren Standardreaktionsenthalpien $\Delta_R H^\ominus$ beliebiger Reaktionen nach $\Delta_R H^\ominus = \sum \nu_i \Delta_B H_i^\ominus$ bestimmt werden. Für die genannten organischen Verbindungen können auf analoge Weise wie die Bildungsenthalpien auch Näherungswerte für die freien molaren Standardbildungs- bzw. Reaktionsenthalpien errechnet werden (↑ Tab. 7.1.2-4).

7.1.2.2.2. *Abschätzung molarer Reaktionsentropien.* Entsprechend der in 6.1.2.3. und 6.1.3.1.2. gegebenen Bestimmung des Entropiebegriffs kann die molare Reaktionsentropie $\Delta_R S$ als Maß für die Änderung der Ordnung bzw. der Bewegungsmöglichkeiten für 1 mol Formelumsätze von den Ausgangsstoffen zu den Endprodukten aufgefaßt werden. Damit lassen sich Vorzeichen und Größenordnung von $\Delta_R S$ für viele Reaktionen abschätzen. $\Delta_R S$ ist im allgemeinen positiv, wenn bei einer Reaktion folgende Fälle vorliegen:

- wenn aus kondensierten Phasen Gase entstehen, beispielsweise $Br_2\ (l) \rightarrow Br_2\ (g)$ mit $\Delta_R S^\ominus = 102\ Jmol^{-1}K^{-1}$;
- wenn reine Stoffe in Lösung gehen, wie $NH_4Cl\ (s) \rightarrow NH_4^+\ (aq) + Cl^-\ (aq)$ mit $\Delta_R S^\ominus = 75{,}2\ Jmol^{-1}K^{-1}$;
- wenn die Zahl der Teilchen erhöht wird, z. B. $2\,H_2O(g) \rightarrow 2\,H_2 + O_2$ mit $\Delta_R S^\ominus = 88{,}6\ Jmol^{-1}K^{-1}$;
- wenn aus starren Molekülen flexiblere entstehen, z. B. Cyclohexan $C_6H_{12} \rightarrow H_3C\text{—}CH_2\text{—}CH_2\text{—}CH_2\text{—}CH\text{=}CH_2$ mit $\Delta_R S^\ominus = 86{,}6\ Jmol^{-1}K^{-1}$.

$\Delta_R S^\ominus$ kann negativ werden, wenn bei der Reaktion kleine oder hochgeladene Ionen gebildet werden, die mit dem polaren Lösungsmittel in starke Wechselwirkung treten und geordnete Solvathüllen ausbilden; z. B. ergibt sich für die Dissoziation des Wassers gemäß $H_2O \rightarrow H^+\ (aq) + OH^-\ (aq)$ $\Delta_R S^\ominus = -80{,}7\ Jmol^{-1}\ K^{-1}$.

Wird $\Delta_R S^\ominus$ aus den molaren Standardentropien der Reaktionsteilnehmer (↑ Tab. 7.1.2-2) bestimmt, dann gilt der Wert jeweils für den Übergang der reinen Ausgangsstoffe in die reinen Produkte. Entsteht aus einer Mischung von Ausgangsstoffen eine Mischung von Produkten, dann ist zusätzlich die Mischungsentropie zu berücksichtigen: $\Delta_R S = \Delta_R S^\ominus + \sum \nu_i[(\Delta_M H_i/T) - R \ln x_i]$, wobei $\Delta_M H_i$ die molare Mischungsenthalpie der Komponenten i darstellt.

7.1.2.3. Grenzen thermodynamischer Aussagen. Die Thermodynamik beantwortet durch Vorzeichen und Betrag von $\Delta_R G$ die Frage, ob und inwieweit die Stoffe eines bestimmten Gemischs miteinander reagieren können. Ist $\Delta_R G$ positiv, dann sind die Ausgangsstoffe stabil in Bezug auf eine Reaktion zu den Endprodukten, ist

Tab. 7.1.2-4 Inkremente I_H bzw. I_G in kJ mol^{-1} zur Berechnung von $\Delta_B H^\ominus$ bzw. $\Delta_B G^\ominus$ gasförmiger organischer Kohlenwasserstoffe bei 298 K; *) nur für Alicyclen anwendbar

Atomgruppe	I_H	I_G	Strukturmerkmal	I_H	I_G
—CH₃	−42,4	−18,0	3 tert. C benachbart	9,6	9,7
—CH₂—	−20,6	8,6	quart. u. tert. C benachbart	10,5	6,8
⟩CH—	−4,6	33,8	2 quart. C benachbart	22,6	10,6
⟩C⟨	3,4	58,5	C_3-Ring*)	101,4	74,4
			C_4-Ring*)	77,0	71,1
			C_5-Ring*)	23,8	−9,4
			C_6-Ring*)	−1,9	−23,4

$\Delta_R G$ negativ, sind sie instabil, und die Produktbildung kann bis zur Gleichgewichtseinstellung ablaufen. Damit ist allerdings noch nichts darüber ausgesagt, ob diese Möglichkeit einer Reaktion tatsächlich auch realisiert wird, ob also die Umsetzung auch mit nennenswerter Geschwindigkeit abläuft. Dieses Problem wird von der chemischen Kinetik behandelt.

Wie entscheidend die Reaktionskinetik für die Antwort auf die Frage ist, ob eine Reaktion tatsächlich abläuft, zeigen die $\Delta_B G^\ominus$-Werte von Tab. 7.1.2-2. Obwohl alle Verbindungen mit positivem $\Delta_B G^\ominus$ instabil sind gegenüber dem Zerfall in die Elemente, sind die wenigsten labil und zerfallen tatsächlich spontan. Die meisten dieser Stoffe, wie z. B. Ethylen $H_2C{=}CH_2$, Blausäure HCN oder die *n*-Alkane oberhalb *n*-Pentan, haben unbeeinflußt eine praktisch unbegrenzte Lebensdauer, sind also *inert*. Existierte diese kinetische Hemmung der Reaktion nicht, würden auch die organischen Verbindungen spontan mit Luftsauerstoff zu Wasser, Kohlendioxid u. a. verbrennen; das würde die Existenz von Leben jeder Art ausschließen.

Der tatsächliche Ablauf der von der Thermodynamik als möglich vorausgesagten Reaktionen mit $\Delta_R G < 0$ ist davon abhängig, ob dem Reaktionssystem die notwendige *Aktivierungsenergie* (↑ 7.2.2., 7.2.3.) zur Bindungslockerung oder -spaltung zugeführt wird. So genügt z. B. ein Funken, um die thermodynamisch erlaubte Verbrennung in einer Ethylen-Sauerstoff-Mischung explosionsartig ablaufen zu lassen, die ansonsten instabil, aber inert existieren kann.

Thermodynamische Aussagen beziehen sich prinzipiell nur auf die Zustände vor und nach der Reaktion; über den Reaktionsmechanismus, also über Details des Wegs von den Ausgangsstoffen zu den Produkten bzw. über die Dynamik der Gleichgewichtseinstellung, wird nichts ausgesagt. Auch hierfür läßt sich ein Bezug zur Kinetik herstellen. Daß die Gleichgewichtseinstellung in Abhängigkeit von den Ausgangsaktivitäten der Partner sowohl durch Reaktion

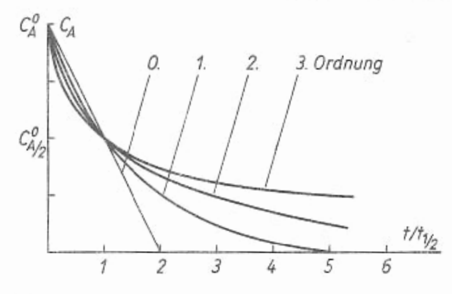


Abb. 7.2.1-1 Zeitverlauf von Reaktionen verschiedener Ordnung bei gleichen Anfangskonzentrationen $c_A{}^\circ$ und gleicher Halbwertszeit

von links nach rechts, also die *Hinreaktion*, als auch durch Reaktion von rechts nach links, d. h. die *Rückreaktion*, erfolgen kann, beweist, daß in jedem Augenblick sowohl elementare Hin- als auch Rückreaktionen ablaufen. Bei der Gleichgewichtseinstellung von links nach rechts ist dann die Geschwindigkeit der Hin- größer als die der Rückreaktion: $v_{\rightarrow} > v_{\leftarrow}$. Ist $v_{\rightarrow} = v_{\leftarrow}$, dann ist das Gleichgewicht erreicht, das folglich den Charakter eines *dynamischen Gleichgewichts* hat.

7.2. Kinetik chemischer Reaktionen

7.2.1. Bestimmung und Aussagen von Zeitgesetzen homogener Reaktionen

7.2.1.1. Reaktionsgeschwindigkeit und Zeitgesetze. Das Fortschreiten einer chemischen Reaktion $|\nu_A|\, A + |\nu_B|\, B + \ldots \rightarrow |\nu_X|\, X + |\nu_Y|\, Y + \ldots$ in einem geschlossenen Raum kann durch die zeitliche Änderung der Konzentrationen von Ausgangsstoffen und Produkten, die *Reaktionsgeschwindigkeit v*, beschrieben werden. Zweckmäßigerweise definiert man diese so, daß sie von der Stöchiometrie der Reaktion unabhängig und auf den Formelumsatz bezogen ist:

$v = (1/\nu_A)\, dc_A/dt = (1/\nu_B)\, dc_B/dt$
$= (1/\nu_X)\, dc_X/dt = (1/\nu_Y)\, dc_Y/dt.$

Für die Reaktion $N_2 + 3\, H_2 \rightarrow 2\, NH_3$ ist demzufolge $v = (1/2)\, dc_{NH_3}/dt = -dc_{N_2}/dt = -(1/3)\, dc_{H_2}/dt$. Reaktionsgeschwindigkeiten sind stets bezüglich der Bruttoreaktion definiert und beschreiben nicht die zeitliche Änderung der Konzentration eventuell auftretender Zwischenprodukte. Bei konstanter Temperatur hängt v nur von den Konzentrationen aller in der Reaktionsmischung vorhandenen Stoffe ab. Der funktionelle Zusammenhang ist das *Zeitgesetz* der Reaktion. Einfache Zeitgesetze haben die Form $v = kc_A^\alpha c_B^\beta c_C^\gamma \ldots$ Die c_i sind die momentanen Konzentrationen der Ausgangsstoffe. Die Summe der Exponenten $\alpha + \beta + \gamma + \ldots$ ist die *Ordnung der Reaktion*. α, β und γ sind die Reaktionsordnungen bezüglich der Stoffe *A*, *B* und *C*. Die Proportionalitätskonstante *k*, die *Geschwindigkeitskonstante* der Reaktion, ist temperaturabhängig.

7.2.1.1.1. *Zeitgesetze einfacher Reaktionen.* Bei Zerfallsreaktionen $A \rightarrow$ Produkte sowie Isomerisierungen ist die Reaktionsgeschwindigkeit häufig proportional der Konzentration des Ausgangsstoffs *A*: $-dc_A/dt = kc_A$. Es liegt ein *Zeitgesetz 1. Ordnung* vor. Beispiele sind der thermische Zerfall von Distickstoffpentoxid $N_2O_5 \rightarrow N_2O_4 + 1/2\, O_2$ oder der radioaktive Zerfall instabiler Atomkerne (↑ 3.3.1.2.). Den zeitlichen Verlauf der Konzentration von *A* erhält man durch Integration des Zeitgesetzes

nach Trennung der Variablen. Mit der Anfangskonzentration c_A^0 bei $t = 0$ ergibt sich: $-\int_{c_A^0}^{c_A} dc_A/c_A = k \int_{t=0}^{t} dt$. Die Lösung lautet $\ln (c_A^0/c_A) = kt$ bzw. $c_A = c_A^0 \, e^{-kt}$. Abb. 7.2.1-1 zeigt den zeitlichen Verlauf von c_A. Die Zeit, nach der die Konzentration des Ausgangsstoffs auf die Hälfte gesunken ist, wird als *Halbwertszeit* $t_{1/2}$ bezeichnet. Für Reaktionen 1. Ordnung gilt $t_{1/2} = (\ln 2)/k = 0{,}693/k$.

Der einfachste Fall eines *Zeitgesetzes 2. Ordnung* (↑ Abb. 7.2.1-1) liegt vor, wenn die Reaktionsgeschwindigkeit dem Quadrat der Konzentration *eines* Ausgangsstoffs proportional ist: $(1/\nu_A)\, dc_A/dt = kc_A^2$. Dieses Zeitgesetz wird z. B. für den thermischen Zerfall von Iodwasserstoff unter Bedingungen gefunden, für die die Rückreaktion vernachlässigt werden kann: $2\,HI \rightarrow H_2 + I_2$. Mit $\nu_A = -2$ und der Anfangskonzentration c_A^0 resultiert nach Trennung der Variablen und Integration für obige Differentialgleichung die Lösung $1/c_A - 1/c_A^0 = 2kt$. Die Halbwertszeit ist $t_{1/2} = 1/2kc_A^0$. Im Unterschied zu Reaktionen 1. Ordnung ist $t_{1/2}$ von der Anfangskonzentration abhängig.

Reaktionen 2. Ordnung $A + B \rightarrow$ Produkte, die jeweils 1. Ordnung bezüglich der Ausgangsstoffe A und B sind, verlaufen nach dem Zeitgesetz $-dc_A/dt = kc_Ac_B$ mit der Lösung $[1/c_B^0 - c_A^0)] \ln (c_A^0 c_B/c_B^0 c_A) = kt$.

Für *Reaktionen n-ter Ordnung* bezüglich eines Ausgangsstoffs gilt das Zeitgesetz $(1/\nu_A)\, dc_A/dt = kc_A^n$. Durch Integration wird die Lösung $[1/(n-1)]\,[1/c_A^{n-1} - 1/(c_A^0)^{n-1}] = |\nu_A|\, kT$ erhalten. Reaktionsordnungen $n > 2$ werden aber relativ selten beobachtet. Nach einem *Zeitgesetz 3. Ordnung* $-(1/2)\; dc_{NO}/dt = kc_{NO}^2 c_{O_2}$ verläuft z. B. die Oxydation von Stickstoffmonoxid $2\,NO + O_2 \rightarrow 2\,NO_2$.

Liegt ein Reaktionspartner in großem Überschuß vor, so daß sich seine Konzentration c im Verlauf der Reaktion praktisch nicht ändert, so kann c mit k zu einer effektiven Konstante k' zusammengefaßt werden. Die Ordnung der Reaktion ist dann anscheinend niedriger und wird als *Pseudoordnung* bezeichnet. Beispiele sind Reaktionen unter Beteiligung eines in großem Überschuß vorliegenden Lösungsmittels oder katalytische Reaktionen (↑ 7.2.4.).

Bei *Reaktionen 0. Ordnung* (↑ Abb. 7.2.1-1) ist die Reaktionsgeschwindigkeit nicht konzentrationsabhängig, sondern wird durch einen zeitlich konstanten nichtchemischen Vorgang bestimmt. Zum Beispiel kann an einer Festkörperoberfläche durch Adsorption aus einem zugeführten Gasstrom die Konzentration des Ausgangsstoffs trotz ablaufender Reaktion konstant gehalten werden.

7.2.1.1.2. *Elementarreaktionen und Reaktionsmolekularität.* Der Ablauf einer Reaktion erfolgt über Zusammenstöße der reagierenden Teilchen. Es ist jedoch offensichtlich, daß eine Reaktion wie die Oxydation von Oxalationen durch Permanganat $2\,MnO_4^- + 5\,C_2O_4^{2-} + 16\,H^+ \rightarrow 2\,Mn^{2+} + 10\,CO_2 + 8\,H_2O$ nicht durch gleichzeitige Wechselwirkung aller beteiligten Ionen erfolgen kann. Sie verläuft wie die meisten chemischen Reaktionen in mehreren Reaktionsschritten, den *Elementarreaktionen.* Diese lassen sich nach ihrer *Molekularität,* d. h. der Zahl der Teilchen, die am zur Reaktion führenden Stoß beteiligt sind, klassifizieren. Bei *monomolekularen Reaktionen* erfolgt Umordnung innerhalb eines Moleküls. *Bimolekulare Reaktionen* erfolgen durch Zusammenstoß von zwei, *trimolekulare Reaktionen* von drei Teilchen. Reaktionen höherer Molekularität sind so unwahrscheinlich, daß sie nicht berücksichtigt werden müssen. Wenn eine Elementarreaktion der Bruttoreaktionsgleichung entspricht, stimmen die Reaktionsordnungen mit den jeweiligen stöchiometrischen Faktoren überein. Bei komplexen Reaktionen wird die Reaktionsgeschwindigkeit durch den am meisten gehemmten Teilschritt bestimmt. Ist dessen Geschwindigkeitskonstante viel kleiner als die der anderen Schritte, ergibt sich eine Reaktionsordnung von 1, 2 oder 3. Verlaufen mehrere Teilschritte mit ähnlicher Geschwindigkeit, werden in der Regel komplizierte Zeitgesetze ermittelt.

Eine experimentell bestimmte ganzzahlige Ordnung der Reaktion sagt also nichts darüber aus, ob eine Elementarreaktion oder eine komplexe Reaktion vorliegt.

7.2.1.2. Experimentelle Bestimmung von Zeitgesetzen

7.2.1.2.1. *Methodik der Messung von Reaktionsgeschwindigkeiten.* Zur Ermittlung der Geschwindigkeitsgleichung einer Reaktion muß die Änderung der Konzentration eines oder mehrerer Reaktionspartner in Abhängigkeit von der Zeit verfolgt werden. Dabei ist die Verwendung äquivalenter Ausgangskonzentrationen vorteilhaft, da in diesem Falle Zeitgesetze der Form $v = kc_A^\alpha c_B^\beta \ldots$ in die einfachere Form $v = kc_A^n$ übergehen. Soll die Reaktionsordnung bezüglich einzelner Komponenten ermittelt werden, setzt man alle anderen Komponenten in so großem Überschuß ein, daß sich ihre Konzentrationen im Verlauf der Reaktion praktisch nicht ändern.

Bei langsam verlaufenden Reaktionen werden die Reaktionspartner in einem temperierbaren Gefäß vermischt und zu bestimmten Zeiten ihre Konzentrationen auf chemischem oder physikalischem Wege ermittelt. Physikalische Verfahren, wie Absorptionsspektroskopie oder Leitfähigkeitsmessungen, haben meist den Vorteil, daß sie schneller als chemische sind und ein Arbeiten direkt im Reaktionsgefäß gestatten.

Zur kinetischen Untersuchung sehr schneller Reaktionen sind *Relaxationsverfahren* geeignet. Man geht vom System im thermodynamischen Gleichgewicht aus und verschiebt durch plötzliche Änderung einer Zustandsvariablen die Gleichgewichtslage. Durch Entladung eines Hochspannungskondensators über die Elektroden einer Meßzelle mit wäßriger Elektrolytlösung können z. B. innerhalb von 10^{-6} s Temperatursprünge von etwa 10 K erzeugt werden. Die Einstellung des neuen Gleichgewichtszustands wird mit sehr schnellen Meßmethoden verfolgt. Bei der *Blitzlichtfotolyse* wird eine Störung des Gleichgewichts durch Einwirkung eines kurzen intensiven Lichtblitzes hervorgerufen. Die Konzentrationsänderungen der kurzlebigen Fotolyseprodukte können mit spektroskopischen Methoden erfaßt werden. Es ist eine Zeitauflösung bis in den Picosekundenbereich möglich.

7.2.1.2.2. *Bestimmung des Zeitgesetzes.* Für Reaktionen 1. Ordnung liefert die Auftragung von $\lg c_A$ über der Zeit eine Gerade, k läßt sich aus dem Anstieg $m = -k/2{,}303$ ermitteln. Bei Zeitgesetzen 2. Ordnung der Form $v = kc_A^2$ liefert die Auftragung von $1/c_A$ über t eine Gerade mit dem Anstieg $2k$ (↑ 7.2.1.1.). Oftmals führen derartige Probierverfahren bereits zum Erfolg. Daneben gibt es zahlreiche Verfahren zur direkten Ermittlung der Reaktionsordnung.

Bei der *Halbwertszeitmethode* wird die Halbwertszeit bei verschiedenen Anfangskonzentrationen bestimmt. Liegt ein Zeitgesetz der Form $v = kc_A^n$ vor, ist die Halbwertszeit gegeben durch $t_{1/2} = (2^{n-1} - 1)/[(n-1)\,|\nu_A|\,k(c_A^0)^{n-1}]$. Nach Logarithmieren ergibt sich $\lg t_{1/2} = \lg\{(2^{n-1}-1)/[(n-1)\,|\nu_A|\,k]\} - (n-1)\lg c_A^0$. Auftragung von $\lg t_{1/2}$ über $\lg c_A^0$ ergibt eine Gerade mit dem Anstieg $m = -(n-1)$. Ist $t_{1/2}$ von der Ausgangskonzentration unabhängig, liegt eine Reaktion 1. Ordnung vor.

Methode der Anfangsgeschwindigkeit. Bei Zeitgesetzen der Form $v = kc_A^\alpha c_B^\beta c_C^\gamma$ läßt sich α ermitteln, indem man zwei Reaktionsverläufe I und II mit gleichem c_B^0 und c_C^0, aber unterschiedlichem c_A^0, untersucht. Für den Anfang der Reaktion gilt $v_{\mathrm{I}}^0/v_{\mathrm{II}}^0 = (c_A^0(\mathrm{I})/c_A^0(\mathrm{II}))^\alpha$, woraus α ermittelt werden kann. Analog erfolgt die Bestimmung von β und γ.

7.2.1.3. Zusammenhang von Kinetik und Reaktionsmechanismus. Ein wesentliches Anliegen kinetischer Untersuchungen ist die Aufklärung des Mechanismus chemischer Reaktionen. Im umfassenden Sinn erfordert das die Unterteilung der Reaktionen in ihre einzelnen Reaktionsschritte, die Charakterisierung der Zwischenverbindungen sowie die vollständige Beschreibung der aktivierten Komplexe (↑ 7.2.2.2.) für jeden Reaktionsschritt bezüglich Bildung, Zusammensetzung, Geometrie, Solvatation, Energie und Zerfall. Da die meisten Reaktionen aus einer mehr oder weniger komplexen Folge von Elementarreaktionen bestehen, existiert eine derartig detaillierte Kenntnis nur für wenige einfache Reaktionen aufgrund quantenchemischer Rechnungen. Allein aus Messungen des makroskopischen zeitlichen Verlaufs der Konzentrationen läßt sich meist kein eindeutiger Zusammenhang zwischen Kinetik und Mechanismus ableiten.

7.2.1.3.1. *Einige Typen zusammengesetzter Reaktionen.* Bei den bisherigen Betrachtungen wurden die Reaktionen als „vollständig" ablaufend angesehen. Häufig darf jedoch die Rückreaktion nicht unberücksichtigt bleiben. Ein Beispiel für derartige *Gleichgewichtsreaktionen* ist das von BODENSTEIN eingehend untersuchte Gleichgewicht $H_2 + I_2 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} 2\,HI$. Bei Annäherung an das Gleichgewicht müssen sowohl die Bildungs- als auch die Zerfallsreaktion berücksichtigt werden. Die Gesamtgeschwindigkeit ist dann $v = 1/2\mathrm{d}c_{HI}/\mathrm{d}t = k_1 c_{H_2} c_{I_2} - k_{-1} c_{HI}^2$. Im Gleichgewicht gilt $v = 0$; das Verhältnis k_1/k_{-1} liefert die Gleichgewichtskonstante K_c. Unabhängig davon, ob von einem äquimolekularen H_2/I_2-Gemisch oder von reinem HI ausgegangen wird, erfolgt eine asymptotische Annäherung an denselben Gleichgewichtszustand (Abb. 7.2.1-2).

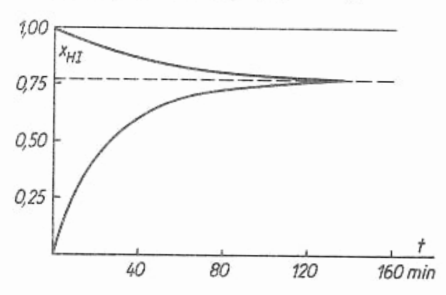


Abb. 7.2.1-2 Molenbruch von HI für die Reaktion $H_2 + I_2 \rightleftharpoons 2\,HI$ als Funktion der Zeit, $\vartheta = 445\,°C$

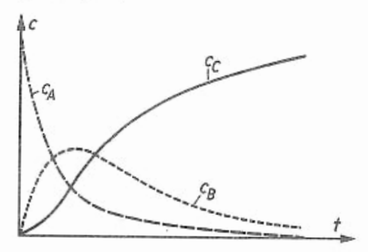


Abb. 7.2.1-3 Konzentrationsverlauf bei einer Folge von zwei Reaktionen 1. Ordnung $A \to B \to C$ als Funktion der Zeit

Folgereaktionen: Entsteht das Endprodukt einer Reaktion über eine Folge von zwei oder mehr Reaktionen 1. Ordnung $A \xrightarrow{k_1} B \xrightarrow{k_2} C$, lauten die Zeitgesetze $dc_A/dt = -k_1c_A$, $dc_B/dt = k_1c_A - k_2c_B$ und $dc_C/dt = k_2c_B$. Nach einer Induktionsperiode, in der vorwiegend B gebildet wird, entsteht zunehmend C. Nach hinreichend langer Zeit zerfällt mehr B als gebildet wird und c_B nähert sich wieder dem Wert Null (Abb. 7.2.1-3).

Häufig werden Folgereaktionen beobachtet, bei denen der erste Schritt reversibel ist. Für die Oxydation von Stickoxid wird z. B. folgender Mechanismus diskutiert: $NO + O_2 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} NO_3$ und $NO_3 + NO \xrightarrow{k_2} 2\,NO_2$. Wegen $k_2 \ll k_{-1}$ ist die erste Teilreaktion stets im Gleichgewicht: $K = k_1/k_{-1} = c_{NO_3}/c_{NO}c_{O_2}$. Für die Geschwindigkeit der Bruttoreaktion folgt $(1/2)\,dc_{NO_2}/dt = k_2c_{NO_3}c_{NO} = k_2Kc_{NO}^2c_{O_2} = k'c_{NO}^2c_{O_2}$, d. h. ein Zeitgesetz 3. Ordnung, wie es auch experimentell gefunden wird.

Bei *Parallel-* oder *Simultanreaktionen* führt die Reaktion eines Ausgangsstoffs A im gleichen Reaktionsgemisch zu verschiedenen Endprodukten: $B \xleftarrow{k_1} A \xrightarrow{k_2} C$.

Verlaufen die Parallelreaktionen nach gleicher Ordnung, verhalten sich die Konzentrationen der Endprodukte wie die entsprechenden Geschwindigkeitskonstanten: $c_B/c_C = k_1/k_2$. Dies gilt allerdings nur bei vollständigem Ablauf der Reaktionen. Bei Systemen $B \underset{k_{-1}}{\overset{k_1}{\leftrightharpoons}} A \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} C$ mit nicht vernachlässigbaren Rückreaktionen hängt das Konzentrationsverhältnis infolge der unterschiedlichen thermodynamischen Stabilität der Produkte von der Reaktionszeit ab. Ist z. B. das Reaktionsprodukt C stabiler, entsprechend $K_1 = k_1/k_{-1} < K_2 = k_2/k_{-2}$, aber $k_1 \gg k_2$, bildet sich zunächst wesentlich mehr B als C. Das System kann jedoch über die mit $\Delta_R G < 0$, d. h. freiwillig ablaufende Reaktion $B \rightarrow A \rightarrow C$ einen stabileren Zustand erreichen. Im Anfangsbereich ist die Reaktion *kinetisch kontrolliert*, der Stoff B überwiegt im Reaktionsgemisch. Hat sich der stabile Endzustand mit hohem Anteil an C eingestellt, ist das Reaktionsergebnis *thermodynamisch kontrolliert*.

Kettenreaktionen sind Folgereaktionen, bei denen durch eine *Startreaktion* energiereiche reaktive Zwischenprodukte gebildet werden, die in Folgereaktionen sehr schnell mit den Ausgangsstoffen reagieren und dabei ständig regeneriert werden. Die *Kettenträger* können durch *Abbruchreaktionen*, z. B. Rekombinationen, verbraucht werden. Kettenträger sind vorwiegend Atome, freie Radikale oder Ionen. Die Zahl der Zyklen bis zum Abbruch der *Reaktionskette* ist die *Kettenlänge*.

Die Synthese von HBr aus den Elementen ist durch einen Kettenmechanismus zu deuten:

Startreaktion: $Br_2 + M \xrightarrow{K_1} 2\,Br + M$

Reaktionskette: $Br + H_2 \xrightarrow{K_2} HBr + H$

$H + Br_2 \xrightarrow{K_3} HBr + Br$

$H + HBr \xrightarrow{K_{-2}} H_2 + Br$

Abbruchreaktion: $2\,Br + M \xrightarrow{K_{-1}} Br_2 + M$.

Start und Abbruch der Reaktionskette bestehen in Zerfall bzw. Bildung eines Br_2-Moleküls unter Mitwirkung eines beliebigen Stoßpartners M. Die Gesamtbildungsgeschwindigkeit von HBr ergibt sich zu $dc_{HBr}/dt = k_2c_{H_2}c_{Br} + k_3c_{Br_2}c_H - k_{-2}c_{HBr}c_H$. Zur Ermittlung der für die Auswertung dieser Gleichung erforderlichen Konzentrationen der Brom- und Wasserstoffatome kann eine von BODENSTEIN eingeführte Näherung, das *Quasistationaritätsprinzip*, verwendet werden. Danach kann die zeitliche Änderung der Konzentration reaktionsfreudiger Zwischenstoffe nach einer Induktionsperiode gleich Null gesetzt werden, da Bildung sowie Zerfall etwa gleich schnell erfolgen und die Konzentrationen der reaktiven Zwischenstoffe immer sehr viel kleiner sind als die der Ausgangs- und Endstoffe. Es ergibt sich in Übereinstimmung mit dem Experiment das Zeitgesetz der HBr-Bildung: $dc_{HBr}/dt = 2k_2(k_1/k_{-1})^{1/2}\,c_{H_2}c_{Br_2}^{1/2}/[1 + (k_{-2}/k_3)\,(c_{HBr}/c_{Br_2})]$.

7.2.1.3.2. *Aufklärung von Reaktionsmechanismen mit Hilfe der Kinetik.* Zur Aufklärung des Mechanismus einer Reaktion ist es zweckmäßig, zunächst aus der Kenntnis von Ausgangs- und Endprodukten und aufgrund der chemischen Erfahrung ein mögliches Reaktionsschema aufzustellen. Dabei braucht zwischen Stöchiometrie der Bruttoreaktion und Reaktionsordnung sowie Mechanismus keinerlei Zusammenhang zu bestehen. Reaktionen gleicher Stöchiometrie können nach völlig unterschiedlichen Mechanismen verlaufen. Mechanismen, die die experimentellen kinetischen Daten nicht richtig widerspiegeln, sind falsch. Demgegenüber ist Übereinstimmung zwischen experimentell ermitteltem und errechnetem Zeitgesetz ein notwendiger, aber kein hinreichender Beweis für die Richtigkeit eines postulierten Mechanismus. Zu seiner Stützung müssen weitere kinetische und nichtkinetische Methoden, wie Isotopenmarkierung, strukturelle Abwandlung von Reaktanten oder Nachweis von Zwischenstoffen, herangezogen werden.

7.2.2. Theoretische Behandlung der Reaktionsgeschwindigkeit

Chemische Reaktion bedeutet im makroskopischen Sinne Wechselwirkung von Stoffen unter Bildung neuer Stoffe und im mikroskopi-

schen Sinne Kontakt der miteinander reagierenden Teilchen, damit die elektronische Umgruppierung vor sich gehen kann. Die Bildung von Komplexen aus den in einem Reaktionsschritt reagierenden Atomen oder Molekülen ist eine Grundvorstellung für die Deutung von Reaktionsabläufen.

Zur Beschreibung der *Temperaturabhängigkeit der Reaktionsgeschwindigkeitskonstante* dient die *Arrheniusgleichung* (1889): $\ln k = \ln A - E_A/(RT)$ oder $k = A \exp[-E_A/(RT)]$ mit der *Aktivierungsenergie* E_A und dem *präexponentiellen Faktor* A. Ein charakteristischer Bestandteil der Arrheniusgleichung ist der *Boltzmannfaktor* $\exp[-E_A/(RT)] = N_{E_A}/N$, der den Bruchteil der Teilchen angibt, deren Energie bei der Temperatur T den Wert E_A übersteigt. Damit wächst k mit dem Anteil der Moleküle, die mit einer Energie $E > E_A$ miteinander wechselwirken. Das ergibt als zweite Grundvorstellung für die Deutung von Reaktionsabläufen, daß Reaktion nur dann eintritt, wenn die Reaktionspartner mit ausreichender Energie im *aktivierten Komplex* zusammentreffen.

Diese qualitativen Vorstellungen führten zu Versuchen, um die Geschwindigkeitskonstante k einer beliebigen Reaktion theoretisch zu berechnen, z. B. mit Hilfe der *Stoßtheorie* (Trautz, Marcelin, Lewis, Herzfeld, Polanyi) sowie der *Theorie der absoluten Reaktionsgeschwindigkeit* oder *des aktivierten Komplexes* bzw. *des Übergangszustands* (Glasstone, Laidler, Eyring, Polanyi).

7.2.2.1. Stoßtheorie. Die Stoßtheorie geht von der Vorstellung aus, daß jede Reaktion durch einen mit ausreichender Energie erfolgenden Zusammenstoß der reagierenden Moleküle, Atome oder Ionen zustande kommt. Dazu wird die Zahl der möglichen Zusammenstöße mit Hilfe der *kinetischen Gastheorie* (↑ 6.2.4.2.1.), die Zahl der davon wirksamen, reaktiven Zusammenstöße mit Hilfe des Boltzmannfaktors berechnet.

7.2.2.1.1. *Bimolekulare Reaktionen.* Ein Zusammenstoß zweier kugelförmiger Teilchen der Radien r_A und r_B und der Massen m_A und m_B erfolgt, wenn sich die Mittelpunkte der Teilchen bis auf den Abstand $(r_A + r_B)$ nähern (Abb. 7.2.2-1). Danach hat ein Teilchen A mit der auf

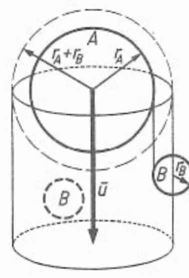


Abb. 7.2.2-1 Zur Stoßtheorie: zylindrische Wirkungssphäre eines Moleküls; Moleküle, deren Mittelpunkt innerhalb des Zylinders liegen, erleiden einen Zusammenstoß

B bezogenen mittleren Relativgeschwindigkeit $\bar{w} = [(8kT/\pi)(1/m_A + 1/m_B)]^{1/2}$ die Wirkungssphäre $\pi(r_A + r_B)^2\,\bar{w}$ und erleidet pro Sekunde $Z' = (r_A + r_B)^2\,\bar{w} N_B$ Zusammenstöße mit den N_B Teilchen pro cm³ von B. Die Anzahl Z_{AB} der Zusammenstöße von N_A Teilchen A pro cm³ mit den N_B Teilchen B beträgt dann pro Sekunde $Z_{AB} = [8\pi kT(1/m_A + 1/m_B)]^{1/2}(r_A + r_B)^2 N_A N_B = Z N_A N_B$, wobei Z die spezifische Stoßzahl für $N_A = N_B = 1$ ist. Würde jeder Stoß zur Umsetzung führen, so wäre die Reaktionsgeschwindigkeit $-\mathrm{d}N_A/\mathrm{d}t = Z_{AB}$ bzw. bei Verwendung der Konzentration c in mol l⁻¹:

$$-\mathrm{d}c_A/\mathrm{d}t = (N_L/1000)\,[8\pi kT(1/m_A + 1/m_B)]^{1/2} \times (r_A + r_B)^2\, c_A c_B.$$

Diese Gleichung stimmt formal mit dem Zeitgesetz $-\mathrm{d}c_A/\mathrm{d}t = k_2 c_A c_B$ einer Reaktion 2. Ordnung überein, wenn man $k_2 = (N_L/1000) \times [8\pi kT(1/m_A + 1/m_B)]^{1/2}(r_A + r_B)^2 = Z$ setzt. Für Z ergibt sich eine Größenordnung von 10^{11} bis 10^{12} l mol⁻¹s⁻¹. Die danach berechneten Geschwindigkeitskonstanten sind aber um einige Zehnerpotenzen größer als die experimentell ermittelten, so daß nicht jeder Zusammenstoß zwischen den Teilchen der Reaktionsteilnehmer erfolgreich sein kann. Ein Reaktionserfolg tritt nur ein, wenn die zusammenstoßenden Teilchen eine bestimmte Mindestenergie, die *Aktivierungsenergie* E, haben. Bei N Partikeln trifft dies nach dem *Boltzmannschen Energieverteilungssatz* jedoch nur auf die Anzahl $N_E = N \exp[-E/(RT)]$ zu. Die berechneten Stoßzahlen sind deshalb mit dem Ausdruck $\exp[-E/(RT)]$ zu multiplizieren, um die Anzahl der erfolgreichen Stöße zu erhalten. Damit lautet das Zeitgesetz der Umsetzung $-\mathrm{d}c_A/\mathrm{d}t = Z c_A c_B \exp[-E/(RT)]$. Für die Geschwindigkeitskonstante ergibt sich daraus $k_2 = Z \exp[-E/(RT)]$.

Die Aktivierungsenergie läßt sich nur in Einzelfällen quantenmechanisch berechnen. In der Regel muß man E zunächst experimentell ermitteln und dann mit Hilfe der Stoßtheorie die Geschwindigkeitskonstante k berechnen. Diese berechneten Werte sind aber häufig immer noch größer als die experimentellen. Man berücksichtigt diese Abweichung durch den empirischen sterischen Faktor p, für den $1 \geqq p > 0{,}1$ (0,01) gilt. Er bringt zum Ausdruck, daß die zusammenstoßenden Partikel eine bestimmte gegenseitige Orientierung haben müssen. Damit ergibt sich für die Geschwindigkeitskonstante insgesamt $k = pZ \exp(-E/RT)$.

Diese Formel entspricht der Arrheniusgleichung (↑ 7.2.2.). Der enthaltene Faktor $A = p(N_L/1000)\,[8\pi RT(1/M_A + 1/M_B)]^{1/2}(r_A + r_B)^2 = pZ$ heißt *Aktionskonstante* oder *Frequenzfaktor* und ist proportional $\sqrt{T}$. Diese Abhängigkeit wurde auch experimentell über einen großen Temperaturbereich beobachtet und fin-

det in der erweiterten Arrheniusgleichung $k = BT^c \exp[-E/(RT)]$ mit $c = 1/2$ Berücksichtigung.

Ein weiterer Grund dafür, daß nicht jeder Stoß zum Reaktionserfolg führen kann, besteht darin, daß die beiden Teilchen A und B die beim Zusammenstoß aufgenommene Energie nicht abführen können und sich deshalb sofort wieder trennen. Damit AB bestehen bleiben kann, muß ein Dreierstoß stattfinden, bei dem ein an der eigentlichen chemischen Reaktion „unbeteiligtes" Teilchen oder die Gefäßwand die Überschußenergie von AB aufnimmt. Da Dreierstöße sehr viel seltener als Zweierstöße sind, kann die Reaktion nur beträchtlich langsamer ablaufen, als der gaskinetischen Stoßzahl entspricht.

7.2.2.1.2. *Monomolekulare Reaktionen.* Echte monomolekulare Reaktionen $A \rightarrow B + C$ sind meist Zerfallsprozesse und verlaufen nach einer Geschwindigkeitsgleichung 1. Ordnung. Bei radioaktiven Zerfallsvorgängen ist die Geschwindigkeitskonstante temperaturunabhängig, bei monomolekularen chemischen Reaktionen befolgt sie hingegen oft die Arrheniusgleichung. Nach LINDEMANN und HINSHELWOOD ist das Auftreten einer Aktivierungsenergie dadurch zu erklären, daß dem eigentlichen Zerfall häufig eine Stoßaktivierung vorausgeht, die zu der Reaktionsfolge $A + A \underset{k_2'}{\overset{k_2}{\rightleftharpoons}} A^* + A$ und $A^* \xrightarrow{k_1} B + C$ führt, in der A^* ein aktiviertes, energiereiches Teilchen A bedeutet; $\mathrm{d}c_B/\mathrm{d}t = k_1 k_2 c_A^2/(k_2' c_A + k_1)$ ist dann die vollständige Geschwindigkeitsgleichung. Bei normalen Drücken ist $k_2' c_A \gg k_1$, und man erhält die Geschwindigkeitsgleichung einer Reaktion 1. Ordnung $\mathrm{d}c_B/\mathrm{d}t = k_1(k_2/k_2')\, c_A$. Bei sehr niedrigen Drücken ist aber umgekehrt $k_1 \gg k_2' c_A$, und die Reaktion verläuft dann nach der 2. Ordnung $\mathrm{d}c_B/\mathrm{d}t = k_2 c_A^2$.

Ein solches Verhalten wird z. B. bei der thermischen Zersetzung von Propionaldehyd CH_3CH_2CHO oder Dimethylether $(CH_3)_2O$ beobachtet.

7.2.2.2. Theorie des aktivierten Komplexes. Die Theorie des aktivierten Komplexes geht davon aus, daß sich im Reaktionsschritt $A + BC \rightarrow AB + C$ die freie Reaktionsenthalpie des Gesamtsystems, ausgehend von den getrennten Ausgangsstoffen über Konfigurationen, in denen die Elektronenumgruppierung zwischen den Reaktionspartnern erfolgt, bis zu den wieder getrennten Produkten, kontinuierlich ändert und dabei durch ein Maximum geht. Die dem Maximum entsprechende Konfiguration der Reaktionsteilnehmer zueinander wird aktivierter Komplex genannt (Abb. 7.2.2-2).

7.2.2.2.1. *Ableitung der Eyringgleichung.* Die Theorie des aktivierten Komplexes geht von der Annahme aus, daß man die Ausgangsstoffe und den aktivierten Komplex als im Gleichgewicht miteinander betrachten und die dafür geltenden thermodynamischen Gesetze anwenden kann. Die Reaktionsgeschwindigkeit wird dann durch den *Zerfall des aktivierten Komplexes* in die Produkte bestimmt.

Die wichtigsten Schritte der Ableitung der Eyringgleichung sind:

Die Reaktionsgeschwindigkeit v ist der Zerfallsfrequenz $\nu^{\ddagger}$, d. h. der Zahl der Zerfallsreaktionen pro Zeiteinheit, proportional: $v = \nu^{\ddagger} c_{(ABC)^{\ddagger}}$.

Die Konzentration des aktivierten Komplexes wird durch das Gleichgewicht $A + BC \rightleftharpoons (ABC)^{\ddagger}$ mit $K^{\ddagger} = c^{\ddagger}_{(ABC)}/(c_A c_{BC})$ bestimmt. Damit ergibt sich das Zeitgesetz $v = \nu^{\ddagger} K^{\ddagger} c_A c_{BC}$, in dem $\nu^{\ddagger} K^{\ddagger} = k$ der Geschwindigkeitskonstante entspricht.

Die Gleichgewichtskonstante $K^{\ddagger}$ wird nach den Gesetzen der statistischen Thermodynamik durch die *Zustandssummen* Q_i (↑ 6.1.3.) der beteiligten Stoffe ausgedrückt:

$K^{\ddagger} = [Q^{\ddagger}_{(ABC)}/(Q_A Q_{BC})] \exp[-E_0/(RT)]$. E_0 ist die Differenz der inneren Energien von Ausgangsstoffen und aktiviertem Komplex bei 0 K. Damit wird die folgende Gleichung erhalten: $k = \nu^{\ddagger} [Q^{\ddagger}_{(ABC)}/(Q_A Q_{BC})] \exp[-E_0/(RT)]$. Um die Zerfallsfrequenz $\nu^{\ddagger}$ zu substituieren, wird von der Zustandssumme des aktivierten Komplexes der Beitrag, der der Schwingung $\nu^{\ddagger}$ in der Zerfallsrichtung entspricht, abgespalten: $Q^{\ddagger}_{(ABC)} = [1 - \exp(-h\nu^{\ddagger}/(k_B T))]^{-1}\, Q^{\ddagger}$. Die Schwingung in der Zerfallsrichtung hat eine sehr geringe Kraftkonstante, da ihre Realisierung doch zum Zerfall in die Produkte führen soll; die Reihenentwicklung für die Exponentialfunktion kann sehr schnell abgebrochen werden: $\exp[-h\nu^{\ddagger}/(k_B T)] \approx 1 - h\nu^{\ddagger}/k_B T + \ldots$. Für die Zustandssumme des aktivierten Komplexes ergibt sich $Q^{\ddagger}_{(ABC)} = [k_B T/(h\nu^{\ddagger})]\, Q^{\ddagger}$. Einsetzen in die Gleichung für k führt zu

$$(1) \quad k = \frac{k_B T}{h} \frac{Q^{\ddagger}}{Q_A Q_{BC}} \mathrm{e}^{-E_0/RT}.$$

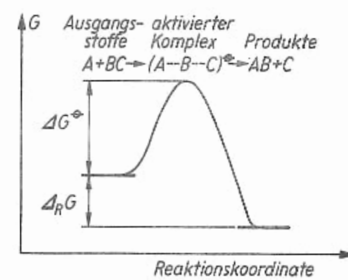


Abb. 7.2.2-2 Änderung der freien Enthalpie G während der Reaktion A + BC → AB + C

Der Ausdruck $[Q^{\ddagger}/(Q_A Q_{BC})] \exp[-E_0/(RT)] = K^{\ddagger\prime}$ entspricht noch näherungsweise der Gleichgewichtskonstanten $K^{\ddagger}$. Mit der aus der Gleichgewichtsthermodynamik entlehnten Beziehung $K^{\ddagger\prime} = \exp[-\Delta G^{\ddagger}/(RT)] = \exp(\Delta S^{\ddagger}/R) \exp[-\Delta H^{\ddagger}/(RT)]$ ergibt sich die Eyringgleichung

$$(2) \quad k = \frac{k_B T}{h} e^{\Delta S^{\ddagger}/R} e^{-\Delta H^{\ddagger}/(RT)}.$$

Die Gleichung 7.2.(1) läßt eine absolute Berechnung der Geschwindigkeitskonstante zu, wenn die Zustandssummen Q_i und die Aktivierungsenergie E_0 beim absoluten Nullpunkt erhältlich sind.

Die Eyringgleichung ist von Bedeutung für die Interpretation experimenteller Ergebnisse in der Sprache der Gleichgewichtsthermodynamik.

7.2.2.2.2. *Potentialflächen.* Sie entstehen durch Auftragen der Energie eines Systems aus Atomen gegen die Koordinaten aller Kerne. Die Energieberechnung kann mit quantenchemischen Methoden unter Ausnutzung der *Born-Oppenheimer-Näherung* (↑ 5.1.2.) erfolgen: Reaktion bedeutet Veränderung der Lage der Atomkerne zueinander, ihre Umgruppierung. Die Kernbewegung erfolgt relativ zur Elektronenbewegung so langsam, daß sich die Elektronen in jedem Augenblick der Kernkonfigurationsveränderung sofort den veränderten Bedingungen anpassen. Deshalb kann die Berechnung $E = f$(Koordinaten, Zeit) durch die punktweise Berechnung $E = f$ (Koordinaten) für alle möglichen Kernkonfigurationen erfolgen. Stabile Kernkonfigurationen machen sich als Mulden auf der Potentialfläche bemerkbar. Der Übergang von einer stabilen Kernkonfiguration in eine andere erfordert immer die Überwindung einer *Potentialschwelle*.

Eine Potentialfläche ist dann noch in der Ebene darstellbar, wenn E die Funktion zweier Koordinaten ist. In diesem Falle werden die Punkte gleicher Energie durch die *Äquipotentiallinien* verbunden und ähnlich den Höhenschichtlinien in den zwei Koordinaten dargestellt.

Sollen in der Reaktion $A + BC \rightarrow AB + C$ nur lineare Kernanordnungen zugelassen sein, dann wird das Reaktionssystem durch die Koordinaten r_{AB} und r_{BC} vollständig beschrieben. Mit BC und AB als jeweils stabile Moleküle gibt Abb. 7.2.2-3 die Darstellung der Potentialfläche in der Ebene wieder. Der Punkt W entspricht der Konfiguration $A \ldots B - C$, bei der BC als stabiles Molekül, unbeeinflußt von A, vorliegt. Entsprechend existiert im Punkt $Y(A - B \ldots C)$ das stabile Molekül AB unbeeinflußt neben C. Der Weg niedrigster Energie(...), die *Reaktionskoordinate*, führt über eine Potentialschwelle, einen Sattelpunkt, von den Ausgangs- zu den Endstoffen. Der *Sattelpunkt* X entspricht dabei dem *aktivierten Komplex* aus den beiden wechselwirkenden Reaktionspartnern, in dem die ursprüngliche Bindung $B - C$ aufgeweitet und die neue Bindung $A - B$ noch nicht vollständig geknüpft ist. Die Auftragung der Energie über der Reaktionskoordinate (↑ Abb. 7.2.2-2), d. h. über dem Minimumweg des Gesamtsystems,

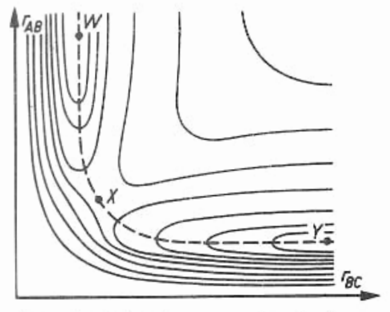


Abb. 7.2.2-3 Höhenschichtendarstellung der potentiellen Energie des Systems A–B–C als Funktion der Kernabstände r_{AB} und r_{BC}

wird *Reaktionsprofil* genannt. Der Energieanstieg bei Annäherung von A an BC rührt von der Abstoßung der sich durchdringenden besetzten Orbitale beider Partner und der erzwungenen Bindungsaufweitung in BC her. Der Energiegewinn nach Durchlaufen des aktivierten Komplexes resultiert aus der Bindungsbildung zwischen A und B.

Ähnliche Potentialflächen werden u. a. für Substitutionsreaktionen der Art $HO^{\ominus} + CH_3{-}Br \rightarrow (HO^{\delta\ominus} \ldots CH_3 \ldots Br^{\delta\ominus})^{\ddagger} \rightarrow HO{-}CH_3 + Br^{\ominus}$ oder Radikalreaktionen der Art: $Cl\cdot + H{-}CH_3 \rightarrow (Cl \ldots H \ldots CH_3)^{\ddagger} \rightarrow Cl{-}H + \cdot CH_3$ erwartet.

Für die Theorie des aktivierten Komplexes liefert die quantenchemisch berechnete Potentialfläche den *Energieunterschied* $\Delta E = E^{\ddagger}_{(ABC)} - E_A - E_{BC}$, der bei Vernachlässigung der Nullpunktschwingungsenergien gleich E_0 in Gleichung 7.2.(1) ist. Außerdem können aus der Potentialfläche die *Geometrie des aktivierten Komplexes* und die *Gestalt der Potentialfläche* um den aktivierten Komplex abgelesen werden, was die Berechnung der Zustandssumme des aktivierten Komplexes $Q^{\ddagger}$ überhaupt erst ermöglicht.

7.2.2.2.3. *Diskussion der Zustandssummen.* Für die Reaktion $A + BC \rightarrow (ABC)^{\ddagger} \rightarrow AB + C$ ergibt sich in Gleichung 7.2.(1) der Quotient der Zustandssummen $Q^{\ddagger}/(Q_A Q_{BC})$. Er kann berechnet werden, wenn die Moleküldaten für A und BC in ausreichender Genauigkeit zur Verfügung stehen und für $(ABC)^{\ddagger}$ einer exakten Potentialflächenberechnung entnommen wer-

den können. Er kann auch für verschiedene Typen von Ausgangsstoffen und Reaktionen aus den Eigenschaften der Zustandssummen abgeschätzt werden.
Ein Molekül kann Energie als *Translations-*, *Rotations-*, *Schwingungs-* und *Elektronenanregungsenergie* speichern (↑ 5.3.). Sie läßt sich jeweils in Anteile für die zur Verfügung stehenden *Freiheitsgrade* aufteilen. Ein nichtlineares Molekül mit n Atomen besitzt 3 Translations-, 3 Rotations- sowie $3n - 6$ Schwingungsfreiheitsgrade; bei einem linearen Molekül sind es 3, 2 bzw. $3n - 5$ Freiheitsgrade. Die Zustandssumme Q_i eines Stoffs läßt sich nun aufschreiben als Produkt der Anteile der verschiedenen Energieformen: $Q_i = Q_{tr}Q_{rot}Q_{vib}Q_{el}$. Jeder Anteil teilt sich in die Anteile pro Freiheitsgrad: $Q_{tr} = f_{tr}^3$; $Q_{rot} = f_{rot}^3$ (bzw. f_{rot}^2) und $Q_{vib} = f_{vib}^{(3n-6)}$ (bzw. $f_{vib}^{(3n-5)}$). Q_{el} ist für Moleküle im elektronischen Grundzustand gleich 1. Für die einzelnen Freiheitsgradtypen ergeben sich die folgenden Größenordnungen der Anteile: $f_{tr} = 10^8$ bis 10^9/cm Molekül; $f_{rot} = 10^1$ bis 10^2/Molekül und $f_{vib} = 10^0$ bis 10^1/Molekül. Für die Reaktion eines Atoms A mit dem zweiatomigen Molekül BC über einen linearen aktivierten Komplex $(ABC)^{\ddagger}$ ergibt sich $Q_A = f_{tr}^3$ und $Q_{BC} = f_{tr}^3 f_{rot}^2 f_{vib}$; für $Q^{\ddagger}$ muß ein Schwingungsfreiheitsgrad weniger gerechnet werden, er wurde bei der Ableitung von Gleichung 7.2.(1) von $Q^{\ddagger}$ abgespalten, so daß dann gilt $Q^{\ddagger} = f_{tr}^3 f_{rot}^2 f_{vib}^3$. Für den Quotienten der Zustandssummen ergibt sich damit: $Q^{\ddagger}/(Q_A Q_{BC}) = f_{vib}^2/f_{tr}^3 \approx 10^{-24}$ cm³/Molekül $\approx 10^0$ cm³mol⁻¹ $\approx 10^{-3}$ lmol⁻¹.
Bei *bimolekularen Reaktionen* ergibt sich in Abhängigkeit von der Größe und Gestalt der Ausgangsstoffe für den Quotienten der Zustandssummen etwa $Q^{\ddagger}/Q_A Q_{BC} \approx 10^{-21}$ bis 10^{-27} cm³/Molekül. Bei *monomolekularen Reaktionen* vereinfacht sich der Quotient zu $Q^{\ddagger}/Q_{BC} = 1/f_{vib} \approx 10^0$ bis 10^1. Diese Abschätzungen stehen in guter Übereinstimmung mit experimentellen Ergebnissen.

7.2.2.2.4. *Vergleich der verschiedenen Gleichungen für die Geschwindigkeitskonstante.* Verglichen werden sollen die Arrheniusgleichung $k = A \exp[-E_A/(RT)]$ (↑ 7.2.2.), die aus der Stoßtheorie abgeleitete Gleichung $k = pZ \exp(-E/RT)$ und die Gleichungen aus der Theorie des aktivierten Komplexes $k = (k_B T/h)[Q^{\ddagger}/(Q_A Q_{BC})] \exp[-E_0/(RT)]$, $k = (k_B T/h) \exp(\Delta S^{\ddagger}/R) \exp[-\Delta H^{\ddagger}/(RT)]$. In allen vier Gleichungen tritt der Boltzmannfaktor mit einer Aktivierungsenergie auf. Aus der Temperaturabhängigkeit der Geschwindigkeitskonstanten (↑ 7.2.3.1.) ergibt sich für den Zusammenhang zwischen den verschiedenen Aktivierungsenergien $E_A = E + 1/2RT = \Delta H^{\ddagger} + RT$. Bei einer üblichen Größenordnung für E_A von 40 bis 140 kJmol⁻¹ kann der Wert $RT = 2{,}5$ kJmol⁻¹ bei Raumtemperatur vernachlässigt werden, und es gilt $E_A \approx E \approx \Delta H^{\ddagger}$.
Damit erweisen sich die übrigen Faktoren als näherungsweise gleich:
$A \approx pZ \approx (k_B T/h)[Q^{\ddagger}/(Q_A Q_{BC})] \approx (k_B T/h) \exp(\Delta S^{\ddagger} R)$.
Da $k_B T/h = 6 \cdot 10^{12}$ s⁻¹ größenordnungsmäßig mit der spezifischen Stoßzahl $Z \approx 10^{11}$ bis 10^{12} lmol⁻¹s⁻¹ übereinstimmt, kann die Beziehung $p \approx Q^{\ddagger}/(Q_A Q_{BC}) \approx \exp(\Delta S^{\ddagger}/R)$ hergestellt werden, d. h., der *sterische Faktor* und der *Quotient der Zustandssummen* repräsentieren den *Entropieeinfluß* auf den Aktivierungsvorgang.

7.2.2.2.5. *Bedeutung der Aktivierungsenthalpie und -entropie.* Die Eyringgleichung (↑ 7.2.2.2.1.) $k = (k_B T/h) \exp(\Delta S^{\ddagger}/R) \exp[-\Delta H^{\ddagger}/(RT)]$ basiert darauf, daß Ausgangsstoffe und aktivierter Komplex miteinander im Gleichgewicht stehen: $A + BC \rightleftharpoons (ABC)^{\ddagger} \rightarrow$ Produkte. Damit können $\Delta H^{\ddagger}$ bzw. $\Delta S^{\ddagger}$ im Sinne der Gleichgewichtsthermodynamik als Änderungen der Standardenthalpie bzw. der Standardentropie beim Übergang von den Ausgangsstoffen zum aktivierten Komplex aufgefaßt und die Überlegungen zu Vorzeichen und Wert von $\Delta_R H^{\ominus}$ bzw. $\Delta_R S^{\ominus}$ (↑ 7.1.2.2.1., 7.1.2.2.2.) auf $\Delta H^{\ddagger}$ und $\Delta S^{\ddagger}$ übertragen werden. Damit wird $\Delta H^{\ddagger}$ proportional zur Differenz der Bindungsenergien für die bei der Bildung des aktivierten Komplexes partiell neugebildeten und gebrochenen Bindungen und daraus folgend für einfache Reaktionen proportional zu $\Delta_R H^{\ominus}$ (↑ 7.1.2.2.1.). So wurden für die Radikalreaktionen in der Gasphase $R-H + X\cdot \rightarrow R\cdot + H-X$ die Regeln $E_A = 0{,}25\Delta_R H^{\ominus} + 48$ kJmol⁻¹ für $\Delta_R H^{\ominus} < 0$ und $E_A = 0{,}75\Delta_R H^{\ominus} + 48$ kJmol⁻¹ für $\Delta_R H^{\ominus} > 0$ von EVANS und POLANYI formuliert. Für $X\cdot = \cdot CH_3$ und $R = CH_3$ wird E_A experimentell zu 53,6 und nach der Regel zu 48,1 kJmol⁻¹ bestimmt. Im Fall $R = (CH_3)_2CH$ ergibt sich für E_A experimentell der Wert 34,7 und nach der Regel 39,7 kJmol⁻¹. Sehr kleine Aktivierungsenergien $\Delta H^{\ddagger} \approx 0$ werden gefunden, wenn in der Reaktion keine Bindungen gespalten, sondern nur neugebildet werden. Das ist der Fall bei Radikalrekombinationen, z. B. $\cdot CH_3 + H\cdot \rightarrow CH_4$, oder Assoziation von Ionen, z. B. $C_6H_5{-}CH_2^+ + Br^- \rightarrow C_6H_5{-}CH_2Br$.
In $\Delta S^{\ddagger}$ drückt sich die Änderung der Bewegungsmöglichkeiten beim Übergang von den Ausgangsstoffen zum aktivierten Komplex aus. *Bimolekulare Reaktionen* sind durch die Wechselwirkung zweier Reaktanten im aktivierten Komplex gekennzeichnet. Der Übergang zum aktivierten Komplex ist daher mit dem Verlust von Translationsfreiheitsgraden verbunden, weshalb $\Delta S^{\ddagger}$ meist < 0 wird. Für die Dimerisation von Butadien in der Gasphase nach

$2\,CH_2{=}CH{-}CH{=}CH_2 \rightarrow$ (Cyclohexen-Ring)$-CH{=}CH_2$

wird z. B. $\Delta S^{\ddagger} = -79{,}5\ \mathrm{Jmol^{-1}K^{-1}}$ gefunden.

Für *monomolekulare Zerfallsreaktionen* bedeutet die Bindungslockerung im aktivierten Komplex eine Erhöhung der Beweglichkeit. Damit ist $\Delta S^{\ddagger} > 0$ zu erwarten. So zeigt der thermische Zerfall von Azo-bis-isobutyronitril in Benzen nach

$NC{-}C(CH_3)_2{-}N{=}N{-}C(CH_3)_2{-}CN \rightarrow 2\,NC{-}C(CH_3)_2{\cdot} + N_2$ ein $\Delta S^{\ddagger} = 43{,}9\ \mathrm{Jmol^{-1}K^{-1}}$.

Bei einem *Zerfall über einen cyclisch aktivierten Komplex*, wie er für den thermischen Zerfall von Estern typisch ist, sind die Bewegungsmöglichkeiten, insbesondere Schwingungen und innere Rotationen, im aktivierten Komplex gegenüber dem Ausgangszustand eingeschränkt, $\Delta S^{\ddagger}$ wird < 0; $\Delta S^{\ddagger} = -21\ \mathrm{Jmol^{-1}K^{-1}}$.

```
CH3 O   CH3
   \ / \ /
    C   CH2
    ‖
    O

     CH3 O          ‡    CH3 O
        \ ⁄ ⋱               \ ⁄⁄
→ (   C   CH2  )  →        C        CH2
      ⁞    ‖               |    +   ‖
      O   CH2              O        CH2
       ⋱ ⁄                  \
        H                    H
```

Bei *polaren Reaktionen in Lösung* wird $\Delta S^{\ddagger}$ sehr stark durch die Änderung der Lösungsmittelordnung beim Übergang zum aktivierten Komplex bestimmt (↑ 7.2.3.3.3.).

7.2.3. Beeinflussung der Geschwindigkeit homogener Reaktionen

Das allgemeine Zeitgesetz für eine chemische Reaktion $(1/\nu_j)\,\mathrm{d}c_j/\mathrm{d}t = k\pi c_i^{n_i}$ besagt, daß die Reaktionsgeschwindigkeit durch die *aktuellen Konzentrationen* der Reaktionsteilnehmer und die *Größe der Reaktionsgeschwindigkeitskonstante* k bestimmt wird (↑ 7.2.1.). k ist eine Funktion der Reaktionsbedingungen, insbesondere von Temperatur T, Druck p und Reaktionsmedium, und ist folglich nur konstant, wenn sich diese *äußeren Einflüsse* während der Reaktion nich ändern. Andererseits kann die Reaktionsgeschwindigkeit durch Variation der äußeren Einflüsse gezielt verändert werden.

Jedes k gilt auch nur für ganz bestimmte Reaktanten und einen definierten Mechanismus. Veränderungen in der Struktur der Reaktanten durch innere, also *elektronische* oder *sterische Einflüsse*, wirken sich gleichfalls auf die Reaktionsgeschwindigkeit aus (↑ 7.3.1.).

7.2.3.1. Temperaturabhängigkeit. Für die meisten chemischen Reaktionen gilt die empirische Regel, daß die Reaktionsgeschwindigkeit bei Temperaturerhöhung um 10 K auf das Zwei- bis Dreifache steigt (Tab. 7.2.3-1). Quantitativ

Tab. 7.2.3-1 Temperaturfaktor $k_{(T+10\mathrm{K})}/k_{(T)}$ und Aktivierungsenergie E_A in $\mathrm{kJmol^{-1}}$ für einige Reaktionen

Reaktion	$k_{(T+10\mathrm{K})}/k_{(T)}$	E_A
PH_3-Zerfall	1,8	42
$CH_3COOC_2H_5 + NaOH$	1,9	46
$KClO_3 + FeSO_4 + H_2SO_4$	2,4	62
$CH_3CH_2ONa + CH_3I$	3,3	85
Rohrzuckerinversion	3,6	91
Enzymwirkung	7	140

ergibt sich aus der Arrheniusgleichung in der Form $\ln k = \ln A - E_A/(RT)$ eine Gerade (Abb. 7.2.3-2) mit dem Anstieg E_A/R und dem Ordinatenabschnitt $\ln A$. E_A, die *Arrheniussche Aktivierungsenergie*, wird aus den experimentellen Geschwindigkeitskonstanten bei zwei verschiedenen Temperaturen ermittelt: $E_A = [2{,}303\,RT_1T_2/(T_1 - T_2)\lg(k_1/k_2)]$. E_A hat als eine dem System zuzuführende Energie ein positives Vorzeichen und liegt meist im Bereich 40 bis $140\ \mathrm{kJmol^{-1}}$ (↑ Tab. 7.2.3-1). A, der *präexponentielle Faktor*, stellt die Geschwindigkeitskonstante bei unendlich hoher Temperatur dar: $A = k(T \rightarrow \infty)$; A kann aus dem Ordinatenabschnitt der Darstellung von $\lg k$ gegen $1/T$ oder durch Einsetzen von E_A und eines $k(T) - T$-Wertepaars in die Arrheniusgleichung bestimmt werden. Mit E_A und A läßt sich unter

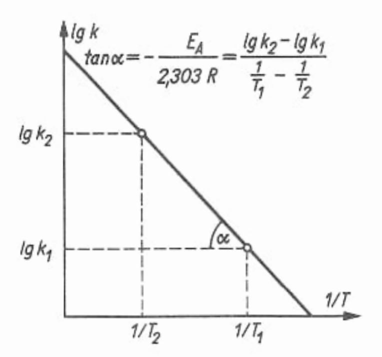


Abb. 7.2.3-2 Zur Berechnung der Aktivierungsenergie E_A aus zwei bei verschiedenen Temperaturen T_1 und T_2 ermittelten Geschwindigkeitskonstanten k_1 und k_2

der Annahme, daß beide Größen in erster Näherung temperaturunabhängig sind, die Geschwindigkeitskonstante bei beliebiger Temperatur angeben.

Damit erweist sich die Temperatur als eine Einflußgröße, die einerseits die *Reaktionsgeschwindigkeit* und andererseits die *Gleichgewichtslage* (↑ 7.1.1.5.2.) stark verändert. Sind sowohl E_A als auch $\Delta_R H^\ominus$ positiv, läuft bei Temperaturerhöhung die Reaktion schneller und zugleich weitgehender zugunsten der Produkte ab. Ist E_A positiv und $\Delta_R H^\ominus$ negativ, stellt sich bei höherer Temperatur das Gleichgewicht schneller ein, ist aber in Richtung der Ausgangsstoffe verschoben. Für die Reaktion $H_2 + I_2 \rightleftharpoons 2\,HI$ ist $\Delta_R H^\ominus = -12{,}36$ kJmol^{-1}, $E_A = 165{,}5$ kJmol^{-1}, $A = 1{,}6 \cdot 10^{17}$ l mol^{-1}s^{-1}. Bei Temperaturanstieg von 125 °C auf 445 °C verringert sich die Gleichgewichtskonstante um den Faktor 0,197 und damit der Reaktionsgrad von 0,89 auf 0,79 (↑ Abb. 7.2.1-2), wenn von $n^\circ_{H_2} = n^\circ_{I_2} = 1$ mol und $n^\circ_{HI} = 0$ ausgegangen wird. Gleichzeitig wird die Geschwindigkeitskonstante der Hinreaktion um den Faktor $4{,}5 \cdot 10^9$ erhöht.

Die Arrheniusgleichung $\ln k = \ln A - E_A/RT$ gilt mit ausreichender Genauigkeit nur für kleine Temperaturintervalle. Über *größere Temperaturbereiche* ergibt die Gleichung $\ln k = \ln B + C \ln T - E/RT$, wie sie mit $C = 1/2$ von der Stoßtheorie abgeleitet wird (↑ 7.2.2.1.), bessere Ergebnisse. Die Theorie des aktivierten Komplexes liefert die Eyringgleichung. $\ln k = \ln (k_B T/h) + \Delta S^\ddagger/R - \Delta H^\ddagger/RT$ mit der Aktivierungsentropie $\Delta S^\ddagger$ und -enthalpie $\Delta H^\ddagger$ (↑ 7.2.2.2.). Für die Temperaturabhängigkeit der Geschwindigkeitskonstante ergibt sich dann $\mathrm{d} \ln k/\mathrm{d}T = E_A/(RT^2) = C/T + E/(RT^2)$ und daraus für den Zusammenhang zwischen den verschiedenen Aktivierungsenergien $E = E_A - CRT$ und $\Delta H^\ddagger = E_A - RT = E + (C - 1)\,RT$. $\Delta H^\ddagger$ kann nach $\lg (k/T) = \lg (k_B/h) + \Delta S^\ddagger/(2{,}303R) - \Delta H^\ddagger/(2{,}303RT)$ aus dem Eyringdiagramm, in dem $\lg (k/T)$ gegen $1/T$ als Gerade aufgetragen ist, als Anstieg $\Delta H^\ddagger/(2{,}303R)$ bestimmt werden. $\Delta S^\ddagger$ ergibt sich aus dem präexponentiellen Faktor nach $\Delta S^\ddagger/(2{,}303R) = \lg (A/T) - \lg \times (\mathrm{e}\, k_B/h)$. Die Aktivierungsgrößen in den Beziehungen für die Temperaturabhängigkeit von k sind nur für enge Bereiche unabhängig von der Temperatur. Vergleiche von $\Delta H^\ddagger$- bzw. $\Delta S^\ddagger$-Werten verschiedener Reaktionen sind deshalb nur zulässig, wenn sie bei ähnlichen Temperaturen gewonnen wurden.

7.2.3.2. Druckabhängigkeit. Genaue experimentelle Untersuchungen zeigen, daß auch in kondensierter Phase die Geschwindigkeiten der chemischen Reaktionen vom Druck p abhängig sind. Zur quantitativen Beschreibung wird das *Aktivierungsvolumen* $\Delta V^\ddagger$ definiert: $(\partial \ln k/\partial p)_T = -\Delta V^\ddagger/RT$. Definitionsgemäß wird $\Delta V^\ddagger$ aus dem Anstieg eines $\lg k$-p-Diagramms gewonnen und erhält ein positives Vorzeichen, wenn durch wachsenden Druck die Geschwindigkeit verringert wird. Experimentell werden bei Reaktionen in kondensierter Phase $\Delta V^\ddagger$-Werte im Bereich 0 bis 20 cm^{3}mol^{-1} gefunden. Das bedeutet aber, daß bei sehr hohen Drücken von mehr als 10^8 Pa (1000 atm) gearbeitet werden muß, um meßbare Effekte zu erhalten. Bei einem $\Delta V^\ddagger = -1$ cm^{3}mol^{-1} z. B. wächst k bei einer Druckerhöhung von 10^5 Pa auf 10^8 Pa lediglich um den Faktor 1,04.

Das Aktivierungsvolumen $\Delta V^\ddagger$ kann im Rahmen der Theorie des aktivierten Komplexes verstanden werden als Differenz zwischen dem partiellen Molvolumen des aktivierten Komplexes $V^\ddagger$ und der Summe der partiellen Molvolumina $\bar{V}_A$, $\bar{V}_B$ der Ausgangsstoffe. Ein positives $\Delta V^\ddagger$ bedeutet danach eine Volumenzunahme beim Übergang zum aktivierten Komplex, z. B. durch Bindungsdehnung; ein negatives $\Delta V^\ddagger$ zeigt an, daß der aktivierte Komplex kompakter ist, z. B. als Folge der Cyclisierung des Ausgangsmoleküls.

Sehr kleine Aktivierungsvolumina $\Delta V^\ddagger \approx 0$ treten bei unpolaren Reaktionen auf, bei denen ein Molekül oder Radikal verbraucht und ein anderes regeneriert wird.

7.2.3.3. Mediumeffekte. Die meisten chemischen Reaktionen laufen in *flüssiger Phase* ab, wo die zwischenmolekularen Wechselwirkungen sehr viel stärker sind als in der Gasphase. Generell wird in Medien höherer Dichte und stärkerer zwischenmolekularer Wechselwirkungen die *Beweglichkeit der Reaktanten* verringert. Deshalb werden beim Übergang von der Gas- zur Flüssigphase oder schon beim Wechsel des Lösungsmittels Veränderungen der Reaktionsgeschwindigkeit bzw. des Reaktionsmechanismus beobachtet.

7.2.3.3.1. *Rolle der Diffusion.* Voraussetzung für den Ablauf einer *bimolekularen Reaktion* ist, daß sich die beiden Reaktanten $A + B$ begegnen. In der Gasphase ist die Diffusionsgeschwindigkeit der Reaktanten zueinander stets größer als die Reaktionsgeschwindigkeit. In Lösung, wo die Diffusion eine Aktivierungsenergie um 12 kJmol^{-1} hat, wird für Reaktionen mit noch kleineren Aktivierungsenergien die Diffusion geschwindigkeitsbestimmend. So sind Radikalrekombinationen und Ionenassoziationen *diffusionskontrolliert*. Ihre Geschwindigkeitskonstanten sind proportional der Summe der Radien und der Diffusionskoeffizienten der Reaktionspartner: $k \sim (r_A + r_B)\,(D_A + D_B)$ und haben die höchsten für Reaktionen in Lösung gefundenen Werte. Sie betragen für Radikalrekombinationen 10^7 bis 10^9 l mol^{-1}s^{-1} und für Ionenassoziationen unter Beteiligung von Protonen sogar bis zu 10^{11} l mol^{-1}s^{-1}.

Entstehen zwei sehr reaktive Partner in Lösung nebeneinander, dann hält bei zu langsamer Diffusion das Lösungsmittel die Partner wie in einem Käfig zusammen; die schnelle Reaktion miteinander ist dann die bevorzugte Alternative. Dieser *Käfigeffekt* wird beobachtet, wenn zwei Radikale bzw. zwei Ionen mit entgegengesetzten Ladungen nebeneinander entstehen. Bei der Fotolyse von Azomethan gemäß $CH_3—N{=}N—CH_3 \xrightarrow[h\nu]{} 2\,CH_3 + N_2$ rekombinieren die beiden Methylradikale zum Ethan H_3CCH_3, ehe sie auseinanderdiffundieren und Folgereaktionen auslösen könnten.

7.2.3.3.2. *Einteilung der Lösungsmittel.* Eine alte Erfahrung besagt, daß sich ein Stoff am besten in dem Lösungsmittel löst, das ihm in Struktur und Eigenschaften am ähnlichsten ist. So ist Toluen $C_6H_5—CH_3$ in Wasser sehr schwer löslich (0,4 g in 100 g H_2O), dagegen mit Benzen C_6H_6, aber auch mit Aceton CH_3COCH_3 und Ethanol CH_3CH_2OH, mischbar. Wasser mischt sich unbegrenzt mit Ethanol.

Meist werden die Lösungsmittel nach ihrer *Polarität* eingeteilt, d. h. nach der Art der *zwischenmolekularen Wechselwirkungen* mit dem gelösten Stoff.

Apolare Lösungsmittel haben niedrige Dielektrizitätskonstanten und Dipolmomente. Bei ihnen dominieren die Dispersionswechselwirkungen. Zu ihnen gehören u. a. die Kohlenwasserstoffe, z. B. Cyclohexan und Benzen, und halogenierte Kohlenwasserstoffe, z. B. Dichlormethan, 1,2-Dichlorethan, Tetrachlormethan.

Dipolar-aprotische Lösungsmittel haben hohe Dielektrizitätskonstanten und Dipolmomente. Sie zeigen neben den Dispersions- zusätzlich beträchtliche Dipolwechselwirkungen. Zur ihnen gehören u. a. Ether, z. B. Diethylether $(CH_3CH_2)_2O$, Aldehyde und Ketone, z. B. Aceton CH_3COCH_3, sowie Dimethylformamid $HCON(CH_3)_2$, Dimethylsulfoxid $(CH_3)_2SO$.

Protische Lösungsmittel haben ein hohes Dipolmoment und können Protonen zur Bildung von *Wasserstoffbrückenbindungen* zur Verfügung stellen. Zu ihnen gehören deshalb u. a. Wasser H_2O, Alkohole, z. B. Ethanol CH_3CH_2OH, und Carbonsäuren, z. B. Essigsäure CH_3COOH.

Als quantitatives Maß der Lösungsmittelpolarität sind die *E_T-Werte* geeignet, deren Zahlenwerte den Energien für den längstwelligen Elektronenübergang eines sehr polaren und stark polarisierbaren Farbstoffs in den Lösungsmitteln entsprechen (Tab. 7.2.3-3).

Der *Donatorcharakter* eines Lösungsmittels wird durch die *Donorzahl DN* widergespiegelt, die die Wechselwirkung des Lösungsmittels mit dem Akzeptor $SbCl_5$ charakterisiert. Die *Akzeptorzahl AN*, die aus der Wechselwirkung des Lösungsmittels mit dem Donator Triethylphosphinoxid, $(C_2H_5)_3P{=}0$, bestimmt wird, kennzeichnet den *Akzeptorcharakter* des Lösungsmittels (↑ Tab. 7.2.3-3).

Die beiden Lösungsmittel Nitromethan und Dimethylsulfoxid haben etwa den gleichen E_T-Wert, unterscheiden sich aber in ihrer spezifischen Wirkung auf das Gelöste. Nitromethan ist ein ausgesprochenes Akzeptorlösungsmittel, während Dimethylsulfoxid sowohl Akzeptor- als auch relativ starken Donatorcharakter besitzt.

7.2.3.3.3. *Lösungsmitteleinfluß auf die Reaktionsgeschwindigkeit.* Er kann bei sehr vielen Reaktionen mit den folgenden Regeln qualitativ vorausgesagt werden:

Zur Abschätzung des Einflusses ist die Wirkung der Lösungsmittelvariation auf die Solvatation der Ausgangsstoffe und des aktivierten Komplexes zu vergleichen.

Beim Übergang von protischen zu dipolar-aprotischen Lösungsmitteln werden Kationen und polare Moleküle stabilisiert und Anionen destabilisiert.

Die Regel von INGOLD und HUGHES besagt: Reaktionen, bei denen im aktivierten Komplex Ladungen gebildet oder konzentriert werden, sollten durch Erhöhung der Polarität des Lösungsmittels beschleunigt werden, während Umsetzungen, in deren aktiviertem Komplex Ladungen verschwinden bzw. delokalisiert werden. gehemmt werden.

Danach sollten unpolare Reaktionen mit einem geschwindigkeitsbestimmenden Schritt der Art: neutrale Ausgangsstoffe → (unpolarer aktivierter Komplex)‡ praktisch unabhängig vom verwendeten Lösungsmittel sein. Der Zerfall des Azo-bis-isobutyronitrils nach

$NC—C(CH_3)_2—N{=}N—C(CH_3)_2—CN$
$\rightarrow N_2 + 2\,NC—C(CH_3)_2\cdot$

folgt dieser Erwartung mit der Geschwindigkeitskonstante $k = 1{,}6 \cdot 10^4\ s^{-1}$ in Toluen und

Tab. 7.2.3-3 E_T-Werte, Akzeptor- und Donorzahlen einiger Lösungsmittel

Lösungsmittel	E_T	AN	DN
Hexan	30,9	0,0	–
CCl_4	32,5	8,6	–
Benzen	34,5	8,2	0,2
Ether	34,6	3,9	19,2
Tetrahydrofuran	37,4	8,0	20,0
$CHCl_3$	39,1	23,1	–
Pyridin	40,2	14,2	33,1
Nitrobenzen	42,0	14,8	4,4
Aceton	42,2	14,8	17,0
Dimethylformamid	43,8	16,0	26,6
Dimethylsulfoxid	45,0	19,3	29,8
Nitromethan	46,3	20,5	2,7
Methanol	55,5	41,3	19,0
Wasser	63,1	54,8	18,0

$k = 1{,}7 \cdot 10^4\ s^{-1}$ in Cyclohexanol. Die Reaktion $CH_3—I + SCN^- \rightarrow CH_3—SCN + I^-$ ist ein Beispiel für den Fall, daß wegen der Delokalisierung der Ladung im aktivierten Komplex $(NCS^{\delta\ominus}...CH_3...I^{\delta\ominus})^{\neq}$ im Vergleich zu den Ausgangsstoffen eine Reaktionshemmung beim Übergang zum polareren Lösungsmittel eintritt. Bei der Reaktion $Br—CH_3 + S(CH_3)_2 \rightarrow (Br^{\delta\ominus}...CH_3...S^{\delta\oplus}(CH_3)_2)^{\neq} \rightarrow Br^- + (CH_3)_3S^+$ werden Ladungen im aktivierten Komplex gebildet, es ist eine Beschleunigung beim Übergang zum polareren Lösungsmittel zu erwarten. Experimentell ergibt sich die Geschwindigkeitskonstante in Dimethylacetamid zu $k = 4 \cdot 10^{-3}\ l\ mol^{-1}s^{-1}$ und in Methanol zu $k = 1 \cdot 10^{-2}\ l\ mol^{-1}s^{-1}$, stimmt damit also überein.

7.2.4. Katalytische Reaktionen in homogenen Medien

Außer durch Temperatur-, Druck- und Lösungsmitteländerungen läßt sich die Geschwindigkeit von Reaktionen durch Anwendung von Katalysatoren beeinflussen. Es handelt sich dabei um Substanzen, die in die Reaktion eingreifen, ohne bleibende Veränderungen zu erfahren. *Positive Katalysatoren* erhöhen die Reaktionsgeschwindigkeit, *negative Katalysatoren* oder *Inhibitoren* vermindern sie. Bei Reaktion eines Ausgangsstoffs über verschiedene Parallelreaktionen zu unterschiedlichen Endprodukten kann die Verwendung eines Katalysators die Vorzugsrichtung der Reaktion dadurch ändern, daß nur eine bestimmte Reaktion beschleunigt wird. Durch eine derartige *selektive Katalyse* lassen sich bei Verwendung unterschiedlicher Katalysatoren aus einem Ausgangsstoff oftmals verschiedene Endprodukte synthetisieren.

7.2.4.1. Wirkungsweise von Katalysatoren. Durch die Mitwirkung eines Katalysators verändert sich der Mechanismus einer Reaktion. Oft bildet der Katalysator mit einem der reagierenden Moleküle einen Zwischenstoff, der wegen seiner anders gearteten Elektronenstruktur zu einer schnellen Weiterreaktion zum Endprodukt unter Rückbildung des Katalysators befähigt ist. Für eine Reaktion $A + B \rightarrow C$ können bei Mitwirkung eines Katalysators folgende Teilschritte formuliert werden: $A + K \rightarrow AK$; $AK + B \rightarrow C + K$. Abb. 7.2.4-1 zeigt die Änderung ΔG der molaren freien Enthalpie für die unkatalysierte und katalysierte Reaktion in Abhängigkeit von der Reaktionskoordinate. Für katalysierte Reaktionen ist eine niedrigere freie Aktivierungsenthalpie der Teilreaktionen charakteristisch, weshalb sie trotz des komplizierteren Mechanismus schneller verlaufen als die entsprechenden unkatalysierten Reaktionen. Da der Katalysator die Energiezustände von Ausgangs- und Endstoffen nicht beeinflußt, bleibt $\Delta_R G$ unverändert, es ist unabhängig vom Reaktionsweg (↑ 7.3.3.3.).

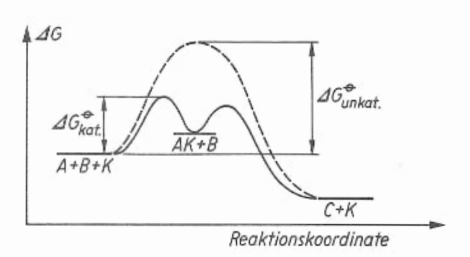


Abb. 7.2.4-1 Molare freie Enthalpie ΔG als Funktion der Reaktionskoordinate; *gestrichelte Kurve* unkatalysierte Reaktion, *durchgezogene Kurve* katalysierte Reaktion

Durch Erhöhung der Geschwindigkeit von Hin- und Rückreaktion beschleunigt ein Katalysator zwar die Einstellung des thermodynamischen Gleichgewichts, auf dessen Lage hat er aber keinen Einfluß.

7.2.4.2. Beispiele homogen-katalytischer Reaktionen

7.2.4.2.1. *Säure-Base-Katalyse.* Eine große Zahl insbesondere organischer Reaktionen wird durch Säuren und Basen katalysiert. Der Katalysator wirkt dabei als *Protonendonator* oder *Protonenakzeptor*. Durch Säure katalysierbar ist z. B. die Verseifung von Carbonsäureestern $RSOOR' + H_2O \xrightarrow{H_3O^+} RSOOH + R'OH$. Der bei Anwesenheit von Säure im Gleichgewicht vorliegende protonierte Ester reagiert viel schneller mit Wasser als der unprotonierte. Nach erfolgter Esterspaltung wird der Katalysator zurückgebildet.

Wird die Esterverseifung in reinem Wasser durchgeführt, so kann die als Reaktionsprodukt entstehende Säure nach teilweiser Dissoziation unter Bildung von H_3O^+-Ionen selbst als Katalysator der Verseifungsreaktion wirken. Eine derartige Erscheinung wird als *Autokatalyse* bezeichnet. Während bei normal katalysierten Reaktionen die Reaktionsgeschwindigkeit infolge Abnahme der Ausgangskonzentrationen ständig zurückgeht, steigt sie bei autokatalytischen Reaktionen durch Bildung des Katalysators zunächst an, um erst gegen Ende der Reaktion, nach weitgehendem Verbrauch der Ausgangsstoffe, abzufallen.

7.2.4.2.2. *Katalyse durch Übergangsmetallkomplexe.* Koordinationsverbindungen von Übergangsmetallen haben wegen ihrer sehr variablen Elektronenstruktur hohe Bedeutung als Katalysatoren. Besonders aktiv sind Metallkomplexe mit labil gebundenen Liganden

oder freien Koordinationsstellen. Als Beispiel soll die Hydroformylierung von Olefinen zu Aldehyden in Gegenwart von Cobalt-Carbonyl-Verbindungen (↑ 5.2.3.4.) dienen:

$$R{-}CH{=}CH_2 + CO + H_2 \xrightarrow[150\,°C]{Katal.} R{-}CH_2{-}CH_2{-}CHO.$$

In Abb. 7.2.4-2 sind die wesentlichen Schritte des Reaktionszyklus dargestellt. Die katalytisch aktiven Spezies sind Verbindungen des Typs $H{-}Co(CO)_3$, an die ein Olefinmolekül unter Ausbildung eines π-Komplexes koordiniert wird (↑ 5.2.3.3.).

7.2.4.2.3. *Enzyme als Katalysatoren.* Im lebenden Organismus ablaufende chemische Reaktionen werden durch *Enzyme* oder *Fermente* katalysiert. Es handelt sich dabei um hochmolekulare Eiweißverbindungen, die im Organismus gebildet werden und als kolloide Partikel vorliegen. Sie bestehen aus einem Eiweißträger, dem *Apoenzym*, und einer *prosthetischen Gruppe*, dem *Coenzym*. Enzyme zeichnen sich durch eine sehr hohe Spezifität ihrer katalytischen Wirkung aus. Nach E. FISCHER paßt das Enzym zum Substrat wie der Schlüssel zum Schloß. Zahlreiche Enzyme weisen auch eine *stereospezifische Wirkung* auf, d. h., sie katalysieren nur die Reaktion eines von zwei möglichen Stereoisomeren. Diese Eigenschaft eröffnet eine Möglichkeit zur *stereospezifischen Synthese* optisch aktiver Verbindungen.

7.2.5. Foto- und strahlenchemische Reaktionen

Vorgänge, bei denen Stoffe und Strahlung unter chemischer Veränderung der Reaktanten in Wechselwirkung treten, sind Gegenstand der Foto- und Strahlenchemie.

Fotochemische Reaktionen werden durch Absorption von Lichtquanten des ultravioletten und sichtbaren Spektralbereichs ausgelöst. Ausgangspunkt sind elektronenangeregte Zustände der Atome, Moleküle oder Ionen. Bei manchen Reaktionen wird die Reaktionsenergie in Form elektromagnetischer Strahlung emittiert. Dieser Vorgang wird als *Chemilumineszenz* bezeichnet. Durch Wechselwirkung mit sehr energiereicher Strahlung wie Röntgen-, γ- oder Elektronenstrahlung werden Stoffe ionisiert; die gebildeten Ionen sind Ausgangspunkt vielfältiger strahlenchemischer Reaktionen.

7.2.5.1. Fotochemie. Zahlreiche chemische Reaktionen verlaufen unter Einwirkung von Licht rasch, im Dunkeln dagegen nur langsam oder gar nicht. Die Dunkelreaktion kann infolge *kinetischer Hemmung* oder aus thermodynamischen Gründen verhindert sein. Nach einer bereits 1838 von GROTTHUS und DRAPER aufgefundenen Gesetzmäßigkeit kann nur die Strahlung fotochemisch wirksam sein, die vom betreffenden Stoff absorbiert wird. Für die Anregung gilt das *fotochemische Äquivalenzgesetz*, wonach für die Anregung *eines* Teilchens gerade *ein* Lichtquant erforderlich ist. Die Geschwindigkeit der fotochemischen Umwandlung eines Stoffs $A \rightarrow B$ ist von der Zahl der absorbierten Lichtquanten abhängig. Ist I_a die Objektmenge der pro Liter und Sekunde absorbierten Lichtquanten, so gilt $dc_A/dt = -\Phi_A I_a$. Es liegt ein Zeitgesetz nullter Ordnung vor (↑ 7.2.1.1.). Φ_A ist die Quantenausbeute der betrachteten Reaktion. Sie ist definiert als Quotient aus der Zahl der umgesetzten Moleküle und der Zahl der absorbierten Lichtquanten. Liegen keine Kettenreaktionen oder Reaktionen angeregter Moleküle mit ihresgleichen im Grundzustand vor, gilt immer $\Phi \leqq 1$. Da angeregte Teilchen A^* in der Regel auf verschiedene Weise weiterreagieren können, ist Φ meist kleiner als eins. Physikalische Folgeprozesse, die zur Desaktivierung führen können, sind z. B. Emission von Strahlung oder strahlungslose Umwandlung der Anregungsenergie in Wärme. Chemische Folgeprozesse sind z. B. Dissoziation, Isomerisierung oder Reaktion mit anderen Atomen oder Molekülen. Die Fotoreaktion ist somit als Parallelreaktion aufzufassen.

Abb. 7.2.4-2 Reaktionszyklus bei der komplexkatalysierten Hydroformylierung eines Olefins

Manche Substanzen L können Teilchen A^* durch Wechselwirkung mit ihnen in den Grundzustand überführen. Ist die Anregungsenergie von A^* ausreichend, um dabei L in den angeregten Zustand zu überführen, kann die Desaktivierung von A^* unter *Energieübertragung* erfolgen, es entsteht ein angeregtes L^*-Molekül. Durch diese *Sensibilisierung* ist es möglich, Stoffe einer fotochemischen Reaktion zugänglich zu machen, die an sich nicht in dem gewünschten Spektralbereich absorbieren.

Die Substanzen L werden als *Löscher* oder *Quencher* bezeichnet, da sie einen zusätzlichen Desaktivierungsweg für A^* eröffnen, durch den

die Wahrscheinlichkeit des Ablaufs der anderen Desaktivierungsprozesse verringert wird.
Ein Beispiel für eine einfache Fotoreaktion vom Typ $A \underset{h\nu}{\overset{h\nu}{\rightleftharpoons}} B$ ist die fotochemische cis-trans-Isomerisierung von Stilben (↑ Abb. 5.2.2-4). Sie ist wie die meisten cis-trans-Isomerisierungen umkehrbar und führt bis zu einem *fotostationären Zustand*. Seine Lage wird durch die Extinktionskoeffizienten ε_A und ε_B bei der Bestrahlungswellenlänge sowie das Verhältnis der Quantenausbeuten für Hin- und Rückreaktion bestimmt. Es gilt die *Zimmermanngleichung* $c_A/c_B = (\Phi_{B \to A}\varepsilon_B)/(\Phi_{A \to B}\varepsilon_A)$.
Bei Stilben liegen im fotostationären Zustand bei einer Einstrahlungswellenlänge von $\lambda =$ 313 nm 93% der Moleküle in cis-Form vor.
Hohe Bedeutung haben Fotoreaktionen für den *fotografischen Prozeß* und die Photosynthese von Kohlenhydraten in den grünen Pflanzen.

7.2.5.2. Strahlenchemie. Die Energiebeträge ionisierender Strahlung sind viel zu groß, als daß sie von der Elektronenhülle der getroffenen Atome oder Moleküle auf einmal aufgenommen werden könnten. Die Energie der eingestrahlten Teilchen wird sukzessive aufgezehrt. Getroffene Atome oder Moleküle werden angeregt bzw. durch Herausschlagen von Elektronen ionisiert. Die entstehenden Produkte sind zu zahlreichen Reaktionen fähig. Meist erfolgen unmittelbar nach ihrer Entstehung weitere Umwandlungen. Durch Rekombination oder Emission von Photonen kann auch eine Rückkehr in den Ausgangszustand erfolgen.
Zur Charakterisierung der Reaktionsausbeute dient in der Strahlenchemie die Energieausbeute *G*. Sie ist definiert als Quotient aus der Zahl umgesetzter Moleküle und der absorbierten Strahlungsenergie in Einheiten von 100 eV.

7.3. Reaktionsverhalten der Stoffe

7.3.1. Reaktionen von Molekülen

7.3.1.1. Klassifizierung chemischer Reaktionen. Unter den zahlreichen individuellen Reaktionen zwischen chemischen Stoffen gibt es Ähnlichkeiten und Verwandtschaften, die zur Klassifizierung der Reaktionen nach sehr verschiedenen Gesichtspunkten benutzt werden. Hierbei sind allerdings anstelle klarer Definitionen oft nur qualitative Formulierungen möglich.

7.3.1.1.1. *Chemische Veränderungen.* Wohl das wichtigste Klassifizierungsmerkmal ist die chemische Veränderung; sie ist in der Bruttoreaktionsgleichung formuliert. Treten zwei Stoffe nach $A + B \to C$ unter Bildung eines neuen Produkts zusammen, spricht man von *Additionen*; z. B. $CH_3{-}CH{=}CH_2 + HCl \to CH_3{-}CHCl{-}CH_3$ oder $SO_3 + H_2O \rightleftharpoons H_2SO_4$.
Die Aufspaltung eines Stoffs in zwei oder mehrere einfache Stoffe nach $A \to B + C$ wird *Zerfalls-* oder *Zersetzungsreaktion* genannt, beispielsweise $CaCO_3 \to CaO + CO_2$ oder $CH_3CH_2{-}O{-}CO{-}CH_3 \to CH_2{=}CH_2 + CH_3COOH$. Behält ein Stoff bei der Reaktion die Zusammensetzung bei und ändert nur seine Struktur, dann liegt eine *Isomerisierung* nach $A \to B$ vor, z. B. $CH_3{-}CH_2{-}CH_2Br \to CH_3{-}CHBR{-}CH_3$. Bei *Austauschreaktionen* werden formal zwischen zwei Reaktionspartnern Fragmente nach $AB + CD \to AC + BD$ ausgetauscht, z. B. $CH_3CH_2{-}OH + HBr \to CH_3CH_2{-}Br + H_2O$ oder $BaS + ZnSO_4 \to BaSO_4 + ZnS$. Bei *Redoxreaktionen* ändern die Reaktionspartner ihre *Oxydationszahlen* (↑ 7.3.1.1.2.).
Viele anorganische Molekülverbindungen können aufgefaßt werden als bestehend aus einem Zentralatom und den dieses umgebenden Liganden, der Ligandensphäre. Dann lassen sich die Reaktionen einteilen in Reaktionen mit *Änderung der Ligandensphäre*, Reaktionen mit *Änderung der Oxydationszahl des Zentralatoms* und *Reaktionen an den Liganden*. Als Änderungen der Ligandensphäre kommen Erhöhungen bzw. Erniedrigungen der Koordinationszahl, d. h. *Additionen* bzw. *Eliminierungen*, in Betracht; z. B.: $F_3B + O(CH_3)_2 \underset{\text{Eliminierung}}{\overset{\text{Addition}}{\rightleftharpoons}} F_3BO(CH_3)_2$. Der Austausch eines Liganden der Ligandensphäre wird als *Substitution* bezeichnet; z. B.: $[Fe(OH_2)_6]^{2+} + NO \rightleftharpoons [Fe(OH_2)_5NO]^{2+} + H_2O$. Außerdem können *Isomerisierungen* der Ligandensphäre auftreten.
Bei der Klassifizierung organisch-chemischer Reaktionen geht man meist von Kohlenstoffatomen als Reaktionszentren aus. Der Ersatz einer Atomgruppe durch eine andere heißt *Substitution*; zum Beispiel: $(CH_3)_3C{-}Cl + H_2O \to (CH_3)_3C{-}OH + HCl$. Hierbei stellt $(CH_3)_3C{-}Cl$ das *Substrat* dar, das durch den Angriff des *Agens* H_2O chemisch verändert wird. Wenn das Agens auch das Lösungsmittel für die Reaktion ist, spricht man von einer *Solvolyse*, im Falle des Wassers von einer *Hydrolyse*. Bei einer *Addition* wird das Agens an das Substrat angelagert, wobei sowohl im Substrat als auch im Agens je eine Bindung gelöst und zwischen beiden zwei neue Bindungen geknüpft werden; z. B.: $C_6H_5{-}CH{=}CH{-}C_6H_5 + Br{-}Br \to C_6H_5{-}CHBr{-}CHBr{-}C_6H_5$. Additionen sind charakteristisch für organische Verbindungen mit Doppel- und Dreifachbindungen. Die Umkehrung der Addition ist die *Eliminierung*. Hierbei werden durch Abspaltung zweier Gruppen von benachbarten Kohlenstoffatomen Mehrfachbindungen, von nicht benachbarten Kohlenstoffatomen Ringverbindungen gebildet; z. B.: $CH_3{-}CH_2{-}OH \to CH_2{=}CH_2 + H_2O$. Wech-

selt in einer Reaktion eine Gruppe im gleichen Molekül von einem Atom zu einem anderen, so ist eine *Umlagerung* abgelaufen; sie ist meist mit Folgereaktionen gekoppelt; zum Beispiel: $(C_6H_5)_2C{=}CHBr \rightarrow C_6H_5{-}C{\equiv}C{-}C_6H_5 + HBr$. Viele Reaktionen organischer Verbindungen laufen unter Änderung der Oxydationszahl des Reaktionszentrums im Substrat ab und können damit auch als *Redoxreaktionen* klassifiziert werden; z. B.: $CH_3{-}\overset{+1}{C}HO + Ag_2O \rightarrow 2\,Ag + CH_3{-}\overset{+3}{C}OOH$. Makromoleküle werden durch *Polymerisation*, *Polyaddition* bzw. *Polykondensation* (↑ 5.4.) synthetisiert.

7.3.1.1.2. *Natur der Reaktionspartner*. Wenn zwei Stoffe miteinander in Wechselwirkung treten, wird eine Reaktion zwischen ihnen besonders dann ablaufen, wenn sie sich in ihrer Valenzelektronenkonfiguration deutlich unterscheiden. Zwei grundlegende Reaktionstypen lassen sich unterscheiden, Reduktions-Oxydations- und Säure-Base-Reaktionen (↑ 7.3.2.6.).

Bei *Redoxreaktionen* tritt ein *Elektronendonator*, d. h. das Reduktionsmittel Red, mit einem *Elektronenakzeptor*, d. h. dem Oxydationsmittel Ox, so in Wechselwirkung, daß eine *Elektronenübertragung* vom Reduktions- zum Oxydationsmittel erfolgt; allgemein: $Red(1) + Ox(2) \leftrightarrows Ox(1) + Red(2)$, z. B.: $Zn + Cu^{2+} \rightleftarrows Zn^{2+} + Cu$ (↑ 7.3.2.6.3.).

Formal kann eine Redoxreaktion in zwei Teilreaktionen aufgeteilt werden:

Bei der *Oxydation* $Red(1) \rightarrow Ox(1) + ne$ werden Elektronen freigesetzt und damit die Oxydationszahl des Stoffs heraufgesetzt, z. B. $Zn \rightarrow Zn^{2+} + 2e$. Durch die mit der Oxydation gekoppelte *Reduktion* $Ox(2) + ne \rightarrow Red(2)$ werden die Elektronen verbraucht und damit die Oxydationszahl des Reaktanten erniedrigt, z. B.: $Cu^{2+} + 2e \rightarrow Cu$.

Die *Oxydationszahl* ist ein nützlicher Hilfsbegriff, der für ein Atom in einem Molekül die Ladung angibt, die es tragen würde, wenn alle seine Bindungen Ionenbeziehungen (↑ 5.1.1.2.) wären.

Für die *Bestimmung der Oxydationszahl* gilt:

Für Elemente ist die Oxydationszahl Null.

Für einfache Ionen ist die Oxydationszahl gleich der Ionenladung.

Die Elektronen, die eine Bindung *A—B* vermitteln, werden vollständig dem Partner mit der größeren Elektronegativität (↑ 5.1.3.3.) zugesprochen.

Bei Bindungen zwischen gleichen Elementen (*A—A*, *B—B*) erhält jeder Partner die Hälfte der Elektronen.

Daraus ergibt sich, daß Fluor als das elektronegativste Element stets die Oxydationszahl −1 aufweist und Metalle im allgemeinen positive Oxydationszahlen haben; Wasserstoff tritt in NH_3 mit der Oxydationszahl +1, in Natriumhydrid mit −1 auf. Da die Änderung der Oxydationszahlen in den Teilreaktionen gleich der Zahl der abgegebenen bzw. aufgenommenen Elektronen ist und für die Bruttoreaktion die Zahl der aufgenommenen und abgegebenen Elektronen gleich sein muß, ist damit eine Vorschrift zur Aufstellung von Redoxgleichungen gegeben. So wird bei der quantitativen Analyse Eisen(II) durch Oxydation mit einer wäßrigen Lösung von Kaliumpermanganat bestimmt. Die entsprechenden Teilreaktionen sind: $Fe^{2+} \rightarrow Fe^{3+} + e$; $\overset{+7}{Mn}O_4^- + 5e + 8\,H^+ \rightarrow Mn^{2+} + 4\,H_2O$. Um die Gesamtreaktion zu erhalten, muß die Oxydation für $5\,Fe^{2+}$ formuliert werden: $5\,Fe^{2+} + MnO_4^- + 8\,H^+ \rightarrow 5\,Fe^{3+} + Mn^{2+} + 4\,H_2O$. Quantitativer Ausdruck der Affinität von einer beliebigen Teilreaktion $Red \rightarrow Ox + n\,e$ ist deren *Elektrodenpotential* (↑ 7.3.4.1.4.).

Die Reaktionspartner in *Säure-Base-Reaktionen* sind jeweils eine Säure und eine Base. Nach Lewis ist eine *Säure* ein *Elektronenpaarakzeptor*, d. h. ein Teilchen mit einer Elektronenlücke, das ein Elektronenpaar aufnehmen kann. Typische Lewissäuren sind danach z. B. Kationen wie Li^+ oder CH_3^+, Moleküle wie BF_3 oder SO_3 mit unvollständigem Elektronenoktett des Zentralatoms, Moleküle wie $TiCl_4$ oder $SnCl_4$, in denen *d*-Orbitale für einen Elektroneneinbau verfügbar sind, aber auch Moleküle, die wegen der innermolekularen Polaritätsverhältnisse Atome mit Elektronendefizit aufweisen, wie z. B. den Carbonyl-Kohlenstoff in Verbindungen $RR'C{=}O$. *Lewisbasen* sind *Elektronenpaardonatoren*, d. h. Stoffe mit verfügbaren Elektronenpaaren. Typische Lewisbasen sind dementsprechend Anionen wie Cl^- oder Methylat CH_3O^-, Moleküle wie $H_2O\rangle$ oder $R_3N|$, die Zentren mit freien Elektronenpaaren besitzen, sowie Substanzen mit polarisierbaren Mehrfachbindungen wie Olefine oder Aromaten.

Die Reaktion zwischen einer Lewissäure und einer Lewisbase nach $S + |B \rightarrow {}^{\ominus}S - B^{\oplus}$ führt zur Ausbildung einer Atombindung und ist der einleitende Schritt vieler Reaktionen.

Beispiel: $(CH_3)_3N| + BBr_3 \rightarrow (CH_3)_3\overset{\oplus}{N}{-}\overset{\ominus}{B}Br_3$.

Eine Weiterentwicklung des Lewis'schen Säure-Base-Konzepts ist das Prinzip der harten und weichen Säuren und Basen nach Pearson (↑ 7.3.1.2.).

Für die Beurteilung der Reaktivität von Verbindungen ist zu beachten, daß jeweils die Unterschiede in der Elektronenstruktur *aller* Reaktionspartner zu berücksichtigen sind. Die Reaktivität einer Verbindung ist somit eine relative Eigenschaft, die entscheidend vom Reaktionspartner geprägt wird.

So wirkt Wasserstoffperoxid, in dem der Sauer

stoff in der Oxydationsstufe −1 vorliegt, gegenüber vielen Stoffen als Oxydationsmittel und wird dabei selbst zu Wasser reduziert, z. B.: $PbS + 4\,H_2O_2 \rightarrow PbSO_4 + 4\,H_2O$. Von starken Oxydationsmitteln wie Kaliumpermanganat wird es dagegen zu Sauerstoff oxydiert, wirkt also als Reduktionsmittel; $2\,KMnO_4 + 5\,H_2O_2 + 3\,H_2SO_4 \rightarrow 2\,MnSO_4 + K_2SO_4 + 5\,O_2 + 8\,H_2O$.

7.3.1.1.3. *Mechanismus.* Die Kenntnis des Mechanismus von Reaktionen (↑ 7.2.1.3.) gestattet deren klare Klassifizierung. Bei komplexen Mechanismen beziehen sich die Aussagen meist auf die *geschwindigkeitsbestimmenden Elementarreaktionen* (↑ 7.2.1.1.2.). Ein wichtiges Merkmal der Reaktionen ist der *Ablauf der Bindungsumgruppierung.* Der Übergang von den Ausgangsstoffen zu den Produkten ist mit Bindungsspaltungen und -knüpfungen verbunden. Erfolgen durch eine kontinuierliche Elektronenumgruppierung zwischen den Reaktionspartnern beide Prozesse gleichzeitig, dann liegt ein *synchroner Reaktionsablauf* vor. Er ist für eine große Zahl von Substitutionen charakteristisch, z. B.: $HO^- + CH_3Br \rightarrow (\overset{\delta\ominus}{HO}\ldots CH_3 \ldots \overset{\delta\ominus}{Br})^{\ominus} \rightarrow HO{-}CH_3 + Br^-$. Der kontinuierliche Umbau der Bindungen während des Reaktionsablaufs hat zur Folge, daß für die Änderung der geometrischen Struktur beim Übergang von den Ausgangsstoffen zu den Produkten die Eigenschaften der beteiligten Orbitale, insbesondere ihre Symmetrie, ausschlaggebend sind. Im obigen Beispiel kann das Hydroxidion OH^- das Substrat nur von der dem Bromatom abgewandelten Seite angreifen. Das bedeutet aber, daß das Kohlenstoffatom als Reaktionszentrum eine Inversion seiner Konfiguration erfährt (Abb. 7.3.1-1).

Reaktionen mit synchronem Ablauf über cyclische aktivierte Komplexe werden als *pericyclische Reaktionen* zusammengefaßt. Ihr Ablauf wird durch das Prinzip der Erhaltung der Orbitalsymmetrie bestimmt. Für die verschiedenen Gruppen pericyclischer Reaktionen wurden die *Woodward-Hoffmann-Regeln* formuliert, die die Reaktionsmöglichkeit der Ausgangsstoffe und die Stereochemie der Produkte beschreiben. Als Beispiel für pericyclische Reaktionen seien die *Diels-Alder-Reaktionen* erwähnt, bei denen durch eine *Cycloaddition* aus einem Dien und einem Olefin ein Sechsring gebildet wird (Abb. 7.3.1-2).

Erfolgen Bindungsbrüche und -bildungen beim Übergang von den Ausgangs- zu den Endprodukten nicht gleichzeitig, liegen *asynchrone Reaktionen* vor. Bei *Substitutionen* z. B. sind zwei mechanistische Alternativen asynchronen Ablaufs möglich. Beim *dissoziativen* oder *Eliminierungs-Additions-Mechanismus* wird zuerst die Bindung zur austretenden Gruppe gelöst, ehe die eintretende Gruppe gebunden wird; z. B.:

$$(CH_3)_3C{-}Cl \rightarrow (CH_3)_3C^+ + Cl^-;$$

$$(CH_3)_3C^+ + OH^- \rightarrow (CH_3)_3C{-}OH.$$

Beim *assoziativen* oder *Additions-Eliminierungs-Mechanismus* liegt die Bindungsneubildung vor der Bindungsspaltung; z. B. bei der Sulfonierung von Aromaten (Abb. 7.3.1-3).

Asynchrone Reaktionen verlaufen im Gegensatz zu synchronen Reaktionen über *reaktive Zwischenstufen*, die sich als kurzlebige Spezies mit physikalischen und chemischen Methoden nachweisen lassen. In organisch-chemischen Reaktionen treten als reaktive Zwischenstufen besonders häufig *Radikale, Carbokationen* und *Carbanionen* auf (↑ 7.3.1.4.).

Sehr oft werden chemische Reaktionen durch die *Molekularität* (↑ 7.2.1.1.2.) des geschwindigkeitsbestimmenden Reaktionsschritts näher ge-

Abb. 7.3.1-1 Stereochemie einer synchronen Substitution

Abb. 7.3.1-2 Beispiel einer Diels-Alder-Reaktion

kennzeichnet. Bei der Substitution bedeutet synchroner Verlauf gleichzeitig eine bimolekulare Reaktion; der asynchron-dissoziative Mechanismus enthält als geschwindigkeitsbestimmenden Schritt die monomolekulare Bindungsspaltung. Trimolekulare Reaktionsschritte fordern das gleichzeitige Zusammentreffen von drei Reaktionspartnern und sind folglich außerordentlich selten.

Reaktion wird deshalb als *elektrophile Substitution am Aromaten*, kurz S_EAr, klassifiziert.
Ist das Agens ein Elektronenpaardonator, d. h. eine Lewisbase, dann reagiert es als Nukleophil bevorzugt mit Partnern, die eine Elektronenlücke bzw. einen Elektronenunterschuß besitzen. Die Substitution in Abb. 7.3.1-1 wird danach als *bimolekulare nukleophile Substitution*, kurz S_N2, bezeichnet.

C_6H_5–H + SO_3 ⟶ σ-Komplex ⟶ C_6H_5–$SO_3^{\ominus}$ + $H^{\oplus}$

Abb. 7.3.1-3 Sulfonierung von Benzen mit konzentrierter Schwefelsäure

Ein weiteres Klassifizierungsmerkmal ist die *Art der Bindungsspaltung* bzw. *-knüpfung*. Behält jeder der bei einer Bindungsspaltung entstehenden Reste ein Elektron des Bindungselektronenpaars, dann wurde die Bindung homolytisch gespaltet, z. B.: $CH_3—CH_3 \rightarrow 2 \cdot CH_3$; die beiden Bruchstücke mit je einem ungepaarten Elektron gehören zu den Radikalen.
Reaktionen, in deren Verlauf Radikale auftreten, werden als *radikalisch* oder *homolytisch* klassifiziert.
Bei einer heterolytischen Bindungsspaltung enthält eines der beiden Bruchstücke das ganze Bindungselektronenpaar und das andere behält eine Elektronenlücke, z. B.: $Li—CH_3 \rightarrow Li^+ + CH_3^-$, d. h., es entstehen eine Lewissäure (Li^+) und eine Lewisbase (CH_3^-). Reaktionen, bei denen Bindungen heterolytisch gespaltet bzw. durch die Wechselwirkung einer Lewissäure mit einer Lewisbase gebildet werden, gehören danach zu den *heterolytischen, polaren* oder *ionischen* Reaktionen.
Die in diesem Abschnitt angeführten Beispiele für Substitutionen sind polare Reaktionen.
Bei vielen Reaktionen wird die chemische Umwandlung durch den Angriff eines kleinen, einfachen Teilchens, des Agens, auf ein größeres, komplizierteres Molekül, das Substrat, eingeleitet. Sehr oft, besonders in der organischen Chemie, zieht man die Natur des Agens zur Klassifizierung der Reaktionen heran. Bei radikalischen Reaktionen ist das Agens ein Radikal, das meist eine Radikalkettenreaktion einleitet (↑ 7.2.1.3.1.). So beginnt bei der Bildung von HBr nach $H_2 + Br_2 \rightarrow 2\,HBr$ die Reaktionskette mit dem Angriff des Bromradikals auf den Wasserstoff H_2 nach $Br\cdot + H_2 \rightarrow HBr + \cdot H$.
In heterolytischen Reaktionen kann das Agens ein Elektronenpaarakzeptor, d. h. eine Lewissäure sein. Es wird folglich einen Partner bevorzugen, der ein Elektronenpaar zur Verfügung stellen kann, d. h. es hat den Charakter eines Elektrophils. In Abb. 7.3.1-3 ist das SO_3 das Elektrophil, das den Aromaten angreift. Die

7.3.1.2. Prinzip der harten und weichen Säuren und Basen. Mit dem *HSAB-Konzept* [engl. Hard and Soft Acids and Bases, HSAB] wird von PEARSON die Lewissche Vorstellung weiterentwickelt, daß die Bildung einer Bindung als Wechselwirkung einer Lewissäure mit einer Lewisbase nach $S^+ + B^- \rightleftharpoons S—B$ beschrieben werden kann. Die Lage des Gleichgewichts hängt ab von der *Ladung der Partner*, die für die elektrostatische Anziehung verantwortlich ist, und von deren *Polarisierbarkeit* (↑ 7.3.4.2.2.), die die Fähigkeit zur Ausbildung von Atombindungen bestimmt (Tab. 7.3.1-4). Für die Wechselwirkung von Säure und Base gilt nach PEARSON: Harte Säuren bevorzugen harte Basen, weiche Säuren bevorzugen weiche Basen.
Die *Addukte* aus harten Säuren und harten Basen werden weitgehend durch *Ionenbeziehungen* zusammengehalten, ihre Bildung ist ladungskontrolliert. Für die Addukte aus weichen Säuren und Basen sind *Atombindungen* charakteristisch; sie entstehen durch Wechselwirkung des höchsten besetzten Molekülorbitals (↑ 5.1.2.) der Basen, aus dem das Bindungselektronenpaar zur Verfügung gestellt wird, mit dem niedrigsten unbesetzten Molekülorbital der Säuren, das durch die Bindungselektronen anteilig genutzt wird. Die Bildung der Addukte erfolgt orbitalkontrolliert.
Das HSAB-Konzept ist geeignet, die *Stabilität von Verbindungen* qualitativ zu erklären. AgI_2^- und I_3^- z. B. sind stabil, weil sie aus den weichen Partnern Ag^+ bzw. I_2 und I^- gebildet werden, während AgF_2^- und I_2F^- nicht existent sind, weil F^- eine harte Base ist. Auch die *Abstufung von Gleichgewichtskonstanten und Reaktionsgeschwindigkeiten* folgt dem HSAB-Konzept. HI ist z. B. eine stärkere Säure als HF, weil die harte Säure H^+ besser mit der harten Base F^- kombiniert als mit der weichen Base I^-. Ebenfalls können mit Hilfe des HSAB-Prinzips *bevorzugte Reaktionsorte* in einem Molekül bestimmt werden. So reagiert z. B. eine weiche Base wie $C_6H_5—S^-$ mit $Cl—CH_2—CH_2—Cl$

unter Substitution an den wenig harten C-Atomen zu $C_6H_5S—CH_2—CH_2—SC_6H_5$, während die harte Base OH^- unter vergleichbaren Bedingungen dem Substrat die harte Säure H^+ unter Bildung von H_2O entreißt und damit eine Eliminierung zu $Cl—CH{=}CH_2 + Cl^-$ einleitet.

Spaltung der C=O-Doppelbindung erfordert etwa den doppelten Energieaufwand wie die Spaltung der C—O-Einfachbindung. Die Bindungsspaltung wird außerdem erleichtert, wenn die entstehenden Radikale durch *Substituenteneffekte* stabilisiert werden. So fordert die Spaltung der $H_3C—H$-Bindung 435 $kJmol^{-1}$, während für die Spaltung der $(C_6H_5)H_2C—H$-Bindung nur 355 $kJmol^{-1}$ aufgebracht werden müssen. Der Ersatz eines H— im Methan durch einen Phenylrest C_6H_5— bewirkt, daß das Radikal $C_6H_5CH_2\cdot$ wesentlich stabiler ist als $H_3C\cdot$, weil das Radikalelektron in die Konjugation des Phenylrings einbezogen wird (↑ 5.1.2.5.). Radikale entstehen aus Molekülen auch dann besonders leicht, wenn die Zerfallsreaktion als Fragmentierung unter gleichzeitiger Bildung eines sehr stabilen Moleküls abläuft. So wird beim Zerfall von Perestern $R—CO—O—OR' \rightarrow R\cdot + \cdot OR' + CO_2$ oder von Azoverbindungen nach $R_3C—N{=}N—CR_3 \rightarrow 2\,R_3C\cdot + N_2$ das thermodynamisch sehr stabile CO_2 bzw. N_2 gebildet.

Tab. 7.3.1-4 Einteilung von Säuren und Basen nach PEARSON

Lewis-Typ	Partialladung am Reaktionszentrum		Polarisierbarkeit	Pearson-Typ		Beispiele (R– = Alkyl–)
Säure	positiv	hoch	gering	Säure	hart	H^+, Be^{2+}, Si^{4+}; BF_3, SO_3
		▽	△		↑ mittel	Zn^{2+}, Sn^{2+}; $B(CH_3)_3$, SO_2, R_3C^+
		gering	hoch		↑ weich	Tl^+, Hg^+, I^+; BH_3, I_2, RS^+
Base	negativ	gering	hoch	Base	weich	CN^-, SCN^-, I^-; R_2S, CO
		△	▽		↓ mittel	Br^-, SO_3^{2-}, NO_2^-, Pyridin, Anilin
		hoch	gering		↓ hart	F^-, OH^-, SO_4^{2-}; H_2O, ROH, R_2O, NH_3, $R—NH_2$

Das HSAB-Konzept ist offensichtlich sehr gut geeignet, polare Wechselwirkungen zwischen zwei Reaktionspartnern zu beschreiben und zu deuten, z. B. den *nukleophilen* bzw. *elektrophilen Angriff* eines Agens auf ein organisches Substrat.

7.3.1.3. Homolyse und Heterolyse kovalenter Bindungen. Bei jeder Reaktion müssen, um neue Stoffe zu bilden, Bindungen gespalten werden. Bei Atombindungen bedeutet das eine Aufteilung des Bindungselektronenpaars auf die ursprünglich verbundenen Atome.

7.3.1.3.1. *Homolyse.* Wird eine Atombindung so gespaltet, daß beide Bruchstücke je ein ungepaartes Elektron behalten, liegt eine Homolyse vor: $A—B \rightarrow A\cdot + B\cdot$, z. B. $H—Cl \rightarrow H\cdot + Cl\cdot$. Die Bruchstücke werden als *Radikale* bezeichnet. Um eine Homolyse zu realisieren, muß dem Molekül die Dissoziationsenergie zugeführt werden, die dem Energiegewinn bei Bildung der Bindung entspricht: $A—B \rightarrow A\cdot + B\cdot$, $\Delta_R H = E_{A-B}$. Für oft auftretende Bindungen sind die *mittleren molaren Bindungsenergien* E_{X-Y} in Tab. 7.1.2-3 zusammengestellt.

Der Energieaufwand zur homolytischen Spaltung einer Bindung hängt von mehreren Faktoren ab. Entscheidend ist die Natur der unmittelbar verbundenen Atome. Besonders Bindungen zu Kohlenstoff und Wasserstoff sind sehr stabil, während die Sauerstoff-Sauerstoff-Einfachbindung, die z. B. in Peroxiden $R—O—O—R$ vorliegt, leicht aufspaltet (↑ Tab. 7.2.3-1). Die Bindungsfestigkeit ist natürlich auch eine Funktion des Bindungsgrads. Die

Die Energie zur Bindungsspaltung wird bei der *Thermolyse* durch Wärmezufuhr aufgebracht. Peroxide und Azoverbindungen werden bereits bei 320 bis 375 K gespaltet, sie dienen deshalb oft als Initiatoren für Radikalreaktionen. C—C- und C—H-Bindungen in organischen Verbindungen brechen erst oberhalb 750 K auf, was z. B. beim Cracken von Erdölfraktionen zur Herstellung von Benzin ausgenutzt wird.

Sehr oft entstehen Radikale durch *Fotolyse*, d. h. durch Einstrahlen von relativ energiereichem Licht. So entspricht blaugrünes Licht der Wellenlänge 500 nm einer Energie von etwa 238 $kJmol^{-1}$, was zur Aufspaltung des Chlors nach $Cl_2 \rightarrow 2\,Cl\cdot$ ausreicht. Die hohe Energie des UV-Anteils im Sonnenlicht vermag viele Radikalreaktionen in der Atmosphäre auszulösen, die u. a. auch zur Bildung des sogenannten Smog führen.

Entsprechend dem relativ hohen Energieaufwand zur Bildung der Radikale sind diese sehr reaktiv und reagieren mit vielen Partnern in Radikalkettenreaktionen (↑ 7.3.1.4.1.).

Der Nachweis der Radikale als reaktive Zwischenstufen komplexer Reaktionen kann mit der sehr empfindlichen Methode der Messung der *paramagnetischen Elektronenresonanz* (↑ 4.4.3.) erfolgen.

7.3.1.3.2. *Heterolyse.* Der Kohlenstoff besitzt eine mittlere Elektronegativität (↑ 5.1.3.3.) und kann dadurch in den organischen Molekülen Bindungen entgegengesetzter Polarität ausbilden. Mit Partnern geringerer Elektronegativität, z. B. Metallen, übernimmt er eine negative Partialladung: $\geq\overset{\delta-}{C}-\overset{\delta+}{M}$. In Bindungen mit elektronegativeren Partnern, z. B. Halogenen, erhält er eine positive Partialladung: $\geq\overset{\delta+}{C}-\overset{\delta-}{Hal}$. Damit werden in Kohlenstoffverbindungen Heterolysen entgegengesetzten Charakters möglich:

Bei einer *kationotropen Heterolyse* verbleibt das Bindungselektronenpaar beim Substrat und der Bindungspartner wird als Kation abgespaltet, z. B. $\geq C-Li \rightarrow \geq C^- + Li^+$.

Bei einer *anionotropen Heterolyse* wird der Bindungspartner mit dem Bindungselektronenpaar als Anion abgespaltet und das Substrat bleibt mit einer Elektronenlücke zurück, z. B.

$\geq C-Cl \rightarrow \geq C^+ + Cl^-$.

Der *Energieaufwand* zur heterolytischen Bindungsspaltung ist in der *Gasphase* stets wesentlich höher als der für die Homolyse, z. B. $CH_3Cl(g) \rightarrow CH_3\cdot(g) + Cl\cdot(g)$, $\Delta_R H = 339$ kJmol^{-1}, aber $CH_3Cl(g) \rightarrow CH_3^+(g) + Cl^-(g)$, $\Delta_R H = 950$ kJmol^{-1}. Deshalb laufen in der Gasphase kaum ionische Reaktionen ab. In *wäßriger Lösung* wird dagegen die Heterolyse wegen der starken Solvatation der Ionen (↑ 7.2.3.3.3.) zur bevorzugten Bindungsspaltung, z. B. $CH_3Cl(aq) \rightarrow CH_3^+(aq) + Cl^-(aq)$, $\Delta_R H = 264$ kJmol^{-1}. Allerdings muß, besonders in unpolaren Lösungsmitteln, auch ein radikalischer Ablauf als Konkurrenzreaktion in Betracht gezogen werden.

Die *Heterolyse* ist in Lösung ein *Zweistufenprozeß.* In der ersten Stufe, der Ionisation, wird die kovalente Bindung gespaltet und das Ionenpaar, das von einer gemeinsamen Solvathülle umgeben ist, gebildet. Im zweiten Schritt, der Dissoziation, entstehen die getrennt solvatisierten freien Ionen.

$$R_3C-Cl_{(solv)} \xrightleftharpoons{\text{Ionisation}} (R_3C^+, Cl^-)_{(solv)} \xrightleftharpoons{\text{Dissoziation}} R_3C^+_{(solv)} + Cl^-_{(solv)}.$$

Die *Ionisation* wird durch Lösungsmittel unterstützt, die die Bindungspolarisierung verstärken, und zwar als Donatoren (↑ 7.2.3.3.2.) durch Wechselwirkung mit dem kationischen bzw. als Akzeptoren (↑ 7.2.3.3.2.) durch Wechselwirkung mit dem anionischen Zentrum.

Die *Dissoziation* wird durch Lösungsmittel sehr hoher Dielektrizitätskonstante ε gefördert.

Im Beispiel $(C_6H_5)_3C-Cl \rightleftharpoons \{(C_6H_5)_3C^+, Cl^-\} \rightleftharpoons (C_6H_5)_3C^+ + Cl^-$ ist die Ionisation im protischen Lösungsmittel Essigsäure CH_3COOH weit fortgeschritten, die Dissoziation wegen des kleinen ε-Werts von 6,2 aber unterdrückt. In Nitrobenzen, das nur einen geringen Akzeptor- und keinen Donatorcharakter hat (↑ Tab. 7.2.3-3), tritt trotz des großen ε-Werts von 34,8 keine Ionisation ein. Damit wird der starke Lösungsmitteleinfluß auf polare Reaktionen verständlich.

Die *Geschwindigkeit* und das *Ausmaß der Heterolyse* einer kovalenten Bindung hängt von der Polarität der Bindung, vom Lösungsmittel und von der Stabilisierung der entstehenden Ionen durch Substituenteneffekte ab.

Substituenten, die Bindungselektronen an sich heranziehen, wie $-NO_2$, $-CN$ oder $-COOR$, stabilisieren Anionen $B|^-$ und destabilisieren Kationen A^+. Substituenten, die Bindungselektronen von sich wegdrücken, wie $-CR_3$, $-O^-$ oder $-COO^-$, wirken umgekehrt.

Das kann am Beispiel der Heterolyse der O—H-Bindung bei der Dissoziation von Carbonsäuren nach $R-CH_2COOH \xrightleftharpoons{K_s} R-CH_2COO^- + H^+$ gezeigt werden. Die Dissoziationskonstante K_s der Nitroessigsäure, $R = NO_2$, $K_s = 2{,}1 \cdot 10^{-2}$ mol l^{-1}, steigt im Vergleich zu der von Essigsäure, $R = H$, $K_s = 1{,}7 \cdot 10^{-5}$ mol l^{-1}, stark an, weil durch die elektronenziehende Nitrogruppe die O—H-Bindung noch stärker polarisiert und das entstehende Anion stabilisiert wird.

Die bei der Heterolyse entstehenden Ionen werden auch stabilisiert, wenn die Ladungszentren in Konjugation zu einem π-Elektronensystem (↑ 5.1.2.5.) stehen, da durch die Einbeziehung der Ladungszentren in das delokalisierte Bindungssystem der Elektronenüberschuß des Anions bzw. das Elektronendefizit des Kations ausgeglichen wird.

Das zeigt wiederum die Heterolyse der O—H-Bindung nach $RO-H \xrightleftharpoons{K_s} RO^- + H^+$ deutlich. Die Dissoziation ist bei Raumtemperatur für Phenol, $R = C_6H_5$, $K_s = 1{,}1 \cdot 10^{-10}$ mol l^{-1} wegen der Einbeziehung des entsprechenden freien Elektronenpaars am Phenolatanion in das π-Elektronensystem des Phenylrings wesentlich stärker als für Wasser, $R = H$, $K_s = 1{,}8 \cdot 10^{-16}$ mol l^{-1}. Die bei der Heterolyse kovalenter Bindungen entstehenden Ionen stellen

im allgemeinen hochreaktive Zwischenstufen dar, die als Lewissäuren bzw. -basen durch elektrophilen bzw. nukleophilen Angriff auf zur Verfügung stehende Substrate die verschiedensten Folgereaktionen einleiten (↑ 7.3.1.4.2.).

7.3.1.4. Reaktive Zwischenstufen. In *asynchronen Reaktionen* (↑ 7.3.1.1.3.) erfolgt die Bindungsumgruppierung beim Übergang von den Ausgangs- zu den Endprodukten schrittweise. Sind die dabei auftretenden Zwischenprodukte hochreaktiv und reagieren sofort mit im Reaktionssystem verfügbaren Partnern weiter, dann werden sie als reaktive Zwischenstufen bezeichnet.

Wegen ihrer hohen Reaktivität bieten sich ihnen meistens mehrere alternative oder parallele Möglichkeiten zur Weiterreaktion. Darin unterscheiden sich die asynchronen wesentlich von den *synchronen Reaktionen.* Als Beispiele für reaktive Zwischenstufen sollen die in der organischen Chemie häufig auftretenden Radikale, Carbeniumionen und Carbanionen, angeführt werden.

7.3.1.4.1. *Radikale.* Ein durch Homolyse (↑ 7.3.1.3.1.) gebildetes Radikal kann unter Erhalt des Radikalcharakters weiterreagieren. Beim *Radikalzerfall* entsteht ein einfacheres Radikal und eine stabile Verbindung, z. B. $(CH_3)_3C—O\cdot \rightarrow \cdot CH_3 + CH_3—CO—CH_3$. Greift es ein nichtradikalisches Substrat unter *Substitution* an, z. B. $Cl\cdot + CH_4 \rightarrow HCl + \cdot CH_3$, entsteht wieder ein Radikal, das seinerseits die Reaktionskette durch eine weitere Substitution, z. B. $\cdot CH_3 + Cl_2 \rightarrow CH_3Cl + Cl\cdot$, fortsetzen kann. Besonders wichtig ist die *Addition* an ungesättigte Verbindungen, zum Beispiel $(CH_3)_3CO\cdot + CH_2{=}CH—Cl \rightarrow (CH_3)_3C—O—CH_2—\dot{C}H—Cl$. Greift das gebildete Radikal immer wieder ein Olefin an, kommt es zur Bildung langer Ketten, zur *Polymerisation,* z. B. $—CH_2—\dot{C}HCl \xrightarrow{+CH_2=CH-Cl} —CH_2—CHCl—CH_2—\dot{C}HCl \xrightarrow{+CH_2=CH-Cl} \ldots$ Auf diese Weise werden viele der technisch bedeutsamen Hochpolymere hergestellt, z. B. PVC oder Piacryl. Durch *Rekombination* der Radikale geht der Radikalcharakter verloren, z. B. $Cl\cdot + \cdot Cl \rightarrow Cl_2$ und $Cl\cdot + \cdot CH_3 \rightarrow CH_3Cl$, die Radikalkettenreaktion bricht ab.

Dementsprechend verlaufen viele Radikalreaktionen nach dem auf die Chlorierung angewendeten Schema:

Kettenstart: $Cl_2 \xrightarrow[h\nu]{} 2\,Cl\cdot$

Reaktionskette: $Cl\cdot + CH_4 \rightarrow HCl + \cdot CH_3$ und $\cdot CH_3 + Cl_2 \rightarrow CH_3Cl + Cl\cdot$

Kettenabbruch: $2\,Cl\cdot \rightarrow Cl_2$; $Cl\cdot + \cdot CH_3 \rightarrow CH_3Cl$ und $2\cdot CH_3 \rightarrow C_2H_6$.

7.3.1.4.2. *Carbeniumionen und Carbanionen.* Für ein Carbeniumion, das durch Heterolyse entstanden ist (↑ 7.3.1.3.2.), z. B. nach $\rangle C—Cl \rightarrow \rangle C^+ + Cl^-$, ist die häufigste Weiterreaktion die *Wechselwirkung mit einem Nukleophil* |B nach $\rangle C^+ + |B \rightarrow \rangle C—B^+$. Ist das Nukleophil das ursprünglich abgespaltene Anion, dann wird die Ausgangsverbindung zurückgebildet. Mit jedem anderen Nukleophil resultiert insgesamt eine *Substitution,* speziell eine *S_N1-Reaktion.* Entstand das Carbeniumion durch elektrophilen Angriff auf eine Doppelbindung, z. B. nach $Br_2 + \rangle C{=}C\langle \rightarrow Br^- + Br—\overset{|}{\underset{|}{C}}—\overset{|}{\underset{|}{C}}{}^+$, dann ist der Angriff des Nukleophils, hier des Br^-, der abschließende Schritt einer *elektrophilen Addition A_E*. Ein Carbeniumion als Elektrophil kann ein vorhandenes Olefin angreifen und damit den kationischen Charakter übertragen nach $\rangle C^+ + CH_2{=}CH—CH_3 \rightarrow —\overset{|}{\underset{|}{C}}—CH_2—\overset{+}{C}H—CH_3$. Durch wiederholten Angriff auf vorhandenes Olefin werden durch *kationische Polymerisation* ebenfalls Hochpolymere erhalten. Die Stabilisierung eines Carbeniumions kann auch dadurch erfolgen, daß ihm eine Lewisbase |B ein Proton vom Nachbar-C-Atom entreißt, wobei eine Doppelbindung entsteht, z. B. $\rangle CH—\overset{|}{\underset{|}{C}}{}^+ + OCH_3^- \rightarrow \rangle C{=}C\langle + CH_3OH$. Zusammen mit der einleitenden Heterolyse liegt damit eine *Eliminierung,* speziell eine *E1-Reaktion,* vor. Hat das Carbeniumion eine ausreichende Lebensdauer, dann wird in bestimmten Fällen auch eine Isomerisierung, z. B. $CH_3—CH_2—CH_2^+ \rightarrow CH_3—\overset{+}{C}H—CH_3$, in das stabilere Ion beobachtet, das dann wie oben weiterreagiert. Generell sind bei Reaktionen über Carbeniumionen Substitutions-, Eliminierungs- und Isomerisierungsprodukte nebeneinander zu erwarten (Abb. 7.3.1-5). Für *Carbanionen* sind die Reaktionsmöglichkeiten analog, allerdings mit entgegengesetzter Polarität.

7.3.1.5. Struktur und Reaktivität von Molekülen. Die Aufklärung dieses Zusammenhangs ist ein Grundanliegen der Chemie. Sie enthält erstens die Frage nach dem für einen bestimmten Angriff bevorzugten Reaktionszentrum in einem Molekül bzw. nach den in einer bestimmten Molekülstruktur erkennbaren Reaktionsmöglichkeiten und zweitens die Frage nach der Beeinflussung einer gegebenen Reaktion durch Strukturvariation.

7.3.1.5.1. *Reaktionsmöglichkeiten eines Moleküls.* Die in einem gegebenen Molekül aufgrund

Abb. 7.3.1-5 Produktverteilung bei der Solvolyse von trans-2-Methyl-cyclohexanoltoluensulfonat in Methanol (OTos-Anion der p-Toluensulfonsäure)

der unterschiedlichen Elektronegativitäten (↑ 5.1.3.3.) der Atome auftretende Polarität und die Polarisierbarkeiten der lokalisierten σ- und π- sowie der delokalisierten π-Bindungen (↑ 5.1.2.) bestimmen den bevorzugten Angriffsort für ein Agens bzw. den Ausgangspunkt für verschiedene Reaktionen.

Ein gutes Beispiel dafür sind Carbonsäuren $R—CH_2—COOH$ und ihre Derivate. Charakteristisches Merkmal dieser Substanzklasse ist ihre Fähigkeit zur *Abspaltung eines Protons* aus der Carboxylgruppe nach $R—CH_2—COOH \rightleftharpoons R—CH_2—COO^- + H^+$ wegen der stark polaren O—H-Bindung und der Stabilisierung des Anions durch Delokalisierung der negativen Ladung über das System $—C{\overset{O}{\underset{O}{\lessgtr}}}^{\ominus}$ unter Einbeziehung der Carbonyldoppelbindung $>C{=}O$. Wegen der hohen Elektronegativität des Sauerstoffs trägt der Kohlenstoff der Carboxylgruppe eine positive Partialladung $—\overset{\delta+}{C}(=\overset{\delta-}{O})—OH$. Ein *nukleophiler Angriff* auf das Carboxyl-C-Atom wird noch durch die leichte Polarisierbarkeit der Carbonylgruppe unterstützt: $B|^- + —C(=O)—OH \longrightarrow (B)(OH)>C—O^-$. Diese Reaktion der Carboxylgruppe liegt z. B. der *Veresterung* nach $CH_3COOH + CH_3OH \rightleftharpoons CH_3CO—O—CH_3 + H_2O$ zugrunde. Durch die elektronenanziehende Wirkung der Carboxylgruppe kommt es weiterhin zu einer *Aktivierung des benachbarten C-Atoms*, wie die leichte Substituierbarkeit einer Gruppe an diesem C-Atom zeigt, z. B. $CH_3—CHBr—COOH + NH_3 \rightarrow CH_3—CH(NH_2)—COOH + HBr$.

7.3.1.5.2. *Reaktivitätsabstufungen.* Die Reaktivität eines Reaktionszentrums in einer bestimmten Reaktion wird sehr stark durch den *elektronischen Einfluß der Substituenten*, die an das Reaktionszentrum gebunden sind, verändert. Elektronenziehende (-drückende) Substituenten unterstützen Reaktionsschritte, bei denen sich die Elektronendichte am Reaktionszentrum erhöht (verringert).

Die elektronische Wirkung der Substituenten kann über das *σ-Bindungssystem* durch Beeinflussung der Polarität der Bindungen am Reaktionszentrum erfolgen. Elektronenziehend (-drückend) wirken Substituenten mit höherer (geringerer) Elektronegativität als das Reaktionszentrum, sie haben einen negativen (positiven) *induktiven Effekt*, kurz *$-I$-($+I$-)Effekt* auf das Reaktionszentrum;

z.B.: $-I$: $—\overset{+}{N}(CH_3)_3 > —NO_2 > —CN > —F > —Cl > —Br > —I > —COOCH_3 > —N(CH_3)_2 > —COCH_3 > —OH > —C_6H_5$;
$+I$: $—C(CH_3)_3 > —CH(CH_3)_2 > —CH_2CH_3 > —CH_3$.

Substituenten mit Mehrfachbindungen bzw. freien Elektronenpaaren, die mit dem Reaktionszentrum in Konjugation stehen, können über das π-Bindungssystem die Elektronendichte am Reaktionszentrum verringern bzw. erhöhen, sie üben einen negativen bzw. positiven *mesomeren Effekt*, kurz *$-M$- bzw. $+M$-Effekt*, aus:

z.B.: $-M$: $—NO_2 > —COCH_3 > —COOCH_3 > —CN$
$+M$: $—NH_2 > —OH > —OCH_3 > —F > —Cl > —Br > —I > —C_6H_5$.

Die Dissoziation der Carbonsäuren $R—CH_2—COOH \rightleftharpoons R—CH_2—COO^- + H^+$ wird nur vom induktiven Effekt des Substituenten R beeinflußt. In Übereinstimmung mit der Erwartung ergibt sich folgende Abstufung der Dissoziationskonstanten K_s bei Raumtemperatur:

$R = H$ $\quad 1{,}7 \cdot 10^{-5}\ mol\ l^{-1}$
$< R = OH$ $\quad 1{,}5 \cdot 10^{-4}\ mol\ l^{-1}$
$< R = NO_2$ $\quad 2{,}1 \cdot 10^{-2}\ mol\ l^{-1}$.

Bei der Dissoziation der Benzencarbonsäuren $R—C_6H_4—COOH \rightleftharpoons R—C_6H_4—COO^- + H^+$ wirken sowohl induktiver als auch mesomerer

Effekt. Es ergibt sich folgende Abstufung der K_s-Werte:
R = OH $2{,}6 \cdot 10^{-5}$ mol l^{-1}
R = H $6{,}3 \cdot 10^{-5}$ mol l^{-1} und
R = NO_2 $3{,}6 \cdot 10^{-4}$ mol l^{-1}.
Für den Fall R = OH überwiegt der $+M$-gegenüber dem $-I$-Effekt und bewirkt dadurch die Veränderung in der Abstufung.

7.3.2. Reaktionen in wäßrigen und wasserähnlichen Lösungen

7.3.2.1. Reaktionstypen in wäßrigen Lösungen. Wasser ist das verbreitetste Lösungsmittel. Es vermag aufgrund seiner hohen Dielektrizitätskonstante und hohen Solvatationstendenz vor allem stark polare Stoffe, z. B. echte und potentielle Elektrolyte, zu lösen. In den entstehenden Elektrolytlösungen (↑ 6.3.4.) stellen sich in Abhängigkeit von Art und Konzentration der gelösten Teilchen, von Lösungsmittelzusammensetzung, Temperatur und anderen Variablen zahlreiche Dissoziations- und Assoziationsgleichgewichte ein, die häufig einander bedingen. Solche Umsetzungen in wäßrigen Lösungen sind ein gutes Beispiel für die in 7.3.1.1.2. getroffene Aussage, daß die Reaktivität von Stoffen immer eine relative, von allen Reaktionspartnern zugleich geprägte Eigenschaft ist. Von den verschiedenen Möglichkeiten der Einteilung von Reaktionen in wäßrigen und wasserähnlichen Lösungen hat sich die durchgesetzt, die den Belangen der analytischen Ausnutzung von Elektrolytgleichgewichten Rechnung trägt. Danach unterscheidet man Säure-Base-, Komplex-, Fällungs- und Redoxgleichgewichte.
Dieser Einteilung liegt die Säure-Base-Theorie von Brönsted und Lowry zugrunde. Durch die Konzepte von Lewis sowie von Pearson (↑ 7.3.1.1.2.) und besonders von Usanovič könnte zwar der Säure-Base-Begriff auf viele, wenn nicht auf alle Elektrolytgleichgewichte ausgedehnt werden, jedoch ist die Säure-Base-Theorie von Brönsted und Lowry wegen ihrer Quantifizierbarkeit vorzuziehen.
Viele quantitative Gesetzmäßigkeiten von Säure-Base-Gleichgewichten lassen sich auch auf andere reversibel und stufenweise verlaufende Ionenkombinationsreaktionen, also Komplex- und Fällungsgleichgewichte, übertragen. Formale Analogien bestehen auch zu Redoxgleichgewichten, obwohl es sich dort um Elektronenaustauschreaktionen handelt. Im Anschluß an die Darstellung von Säure-Base-Gleichgewichten werden deshalb in 7.3.2.6. die übrigen Elektrolytgleichgewichte nur kurz abgehandelt, auf Spezialfälle wie Extraktions- und Ionenaustauschgleichgewichte wird ganz verzichtet. Zu den Redoxreaktionen in homogener Phase wird sich durch die Behandlung elektrochemischer Reaktionen (↑ 7.3.4.), also heterogener Gleichgewichte mit geladenen Teilchen, ein breiterer Zugang ergeben.

7.3.2.2. Säure-Base-Konzept (nach Brönsted und Lowry). Hier werden Stoffe nach der Fähigkeit bewertet, Protonen abzugeben oder aufzunehmen: Säuren sind *Protonendonatoren*, Basen sind *Protonenakzeptoren*. Die Ladung der an Brönstedgleichgewichten beteiligten Teilchen ist für deren Säure- oder Basenatur unwesentlich; es gibt Neutral-, Kation- und Anionsäuren sowie die entsprechenden Basen.
Jede Säure geht durch Abgabe eines Protons in die korrespondierende oder konjugierte Base entsprechend veränderter Ladung über. In kondensierter Phase sind freie Protonen nicht existent, durch Reaktion mit Lösungsmittelmolekülen entstehen daraus solvatisierte Teilchen. *Hydroniumionen* H_3O^+ wurden in den kristallinen Monohydraten einiger starker Säuren wie Perchlorsäure $HClO_4$, Salpetersäure HNO_3 und Salzsäure mit Sicherheit nachgewiesen. In verdünnt wäßrigen Lösungen verbrücken H_3O^+-Ionen weitere Wassermoleküle zu Aggregaten $[H(OH_2)_n]^+$, die rasch ihr Proton an benachbarte Wassermoleküle weitergeben können.
Die Solvatation von Protonen erfolgt unter beträchtlichem Energiegewinn, wodurch die Abspaltung von Protonen aus wasserstoffhaltigen Verbindungen begünstigt wird. Ein korrespondierendes *Säure-Base-Paar* wird also erst wirksam, wenn eine weitere Brönstedbase zugegen ist, die die abgespalteten Protonen aufnimmt und damit selbst in ihre korrespondierende Säure übergeht. Das sich einstellende *Protolysegleichgewicht* $S_1 + B_2 \rightleftharpoons B_1 + S_2$ wird in seiner Lage durch die relative Stärke der beteiligten Brönstedsäuren und -basen bestimmt. Oft wird die Base B_2 vom Lösungsmittel bereitgestellt, z. B. gemäß $CH_3COOH + H_2O \rightleftharpoons H_3O^+ + CH_3COO^-$. Gerade in der Möglichkeit, das Lösungsmittel zwanglos in die Betrachtungen einbeziehen zu können, liegt eine Stärke des Brönstedkonzepts.
Der Zusammenhang zwischen der Struktur eines Stoffs und seinen Säure-Base-Eigenschaften kann folglich immer nur in bezug auf ein bestimmtes anderes Säure-Base-Paar, z. B. auf $H_3O^+ \rightleftharpoons H_2O + H^+$ oder $H_2O \rightleftharpoons OH^- + H^+$, angegegeben werden. Weder Essigsäure Hac noch Ammoniak oder Wasser lassen sich eindeutig als Säure oder Base ansprechen. Gegenüber stärkeren Protonendonatoren wirken diese Stoffe als Brönstedbasen und gehen in die konjugierten Säuren H_2ac^+, NH_4^+ und H_3O^+ über, an stärkere Protonenakzeptoren verlieren sie ihr Proton unter Bildung von ac^-, NH_2^- und OH^-.
Liegen solche amphoteren Substanzen oder *Ampholyte* [*amphoteroi* griech., beide; *lyein* griech., lösen] in reiner Form vor, gehen sie in

geringem Ausmaß mit sich selbst eine Säure-Base-Reaktion ein, die als *Autoprotolyse* bezeichnet wird: $HY + HY \rightleftharpoons H_2Y^+ + Y^-$ (Y = OH, NH_2, ac). Diese Eigendissoziation kann durch Messung der elektrischen Leitfähigkeit der sorgfältig gereinigten Lösungsmittel quantitativ erfaßt werden (↑ Tab. 7.3.2-6). Abb. 7.3.2-1 zeigt, daß für Wasser das *Ionenprodukt*

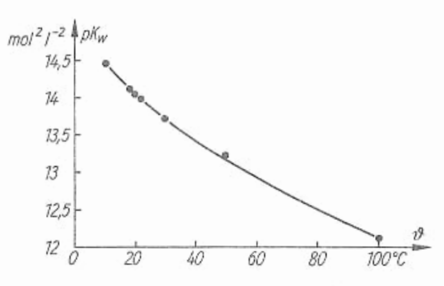


Abb. 7.3.2-1 Temperaturabhängigkeit der Autoprotolysekonstante von Wasser; K_W Ionenprodukt in mol² l⁻²

des Wassers, $K_w = a_{H_3O^+} \cdot a_{OH^-}/a^2_{H_2O}$ zwar stark von der Temperatur abhängig, generell aber nur sehr klein ist. Man kann deshalb die Aktivität des undissoziierten Wassers $a_{H_2O} = 1$ setzen, und außerdem von den Aktivitäten $a_{H_3O^+}$ und a_{OH^-} zu den Ionenkonzentrationen $[H_3O^+]$ und $[OH^-]$ übergehen. Übrigens wird diese Vereinfachung oft noch bei wesentlich höheren Ionenkonzentrationen bis etwa 0,1 mol l^{-1} vorgenommen, wenn auch in weniger guter Näherung. Zur Vermeidung der schwerfälligen Potenzschreibweise gibt man für K_w, $a_{H_3O^+}$, a_{OH^-} und die entsprechenden Konzentrationsausdrücke K^c_w, $[H_3O^+]$ und $[OH^-]$ die negativen dekadischen Logarithmen an. Somit ergibt sich: $pK_w = p_aH + p_aOH$ bzw. $pK^c_w = p_cH + p_cOH$. Neutrales Wasser, in dem die beiden Ionen $[OH^-]$ und $[H_3O^+]$ in gleichen Konzentrationen vorliegen, hat bei verschiedenen Temperaturen recht verschiedene *p*H-Werte (↑ Abb. 7.3.2-1); nur bei Raumtemperatur ergibt sich *p*H = 7,0. Liegen in wäßrigen Lösungen mehr Hydroxid- als Hydroniumionen vor, steigt *p*H über den Neutralpunktswert, die Lösung wird als *basisch* oder *alkalisch* bezeichnet; *sauer* sind Lösungen, deren *p*H-Wert unter dem Neutralwert liegt.
Durch Zugabe entsprechender Protonendonatoren oder -akzeptoren, z. B. H_2SO_4, CH_3COOH, HCO_3^-, HCN, NH_3, OH^-, kann der *p*H- bzw. *p*OH-Wert wäßriger Lösungen in weiten Grenzen – ohne Änderung des Aktivitätsionenprodukts K_w – variiert werden. Eine scharfe Grenze der *p*H- bzw. *p*OH-Skala läßt sich nicht angeben. Bei einmolaren oder noch konzentrierteren Lösungen der entsprechenden Brönstedsäuren und -basen widerspiegeln die Konzentrationsangaben die Aktivitäten von $a_{H_3O^+}$ und a_{OH^-} auch nicht annähernd, nehmen die Aktivitätskoeffizienten widersprüchliche Werte an und komplizieren sich die *p*H-Messungen. Die Angabe, daß die *p*H-Werte wäßriger Lösungen zwischen 0 und 14 liegen, ist insofern eine starke Vereinfachung.

7.3.2.3. Säure- oder Basenstärke in wäßriger Lösung. Verallgemeinert man den Sonderfall der Autoprotolyse wieder auf Umsetzungen zwischen zwei verschiedenen Säure-Base-Paaren, z. B. $HA + H_2O \rightleftharpoons A^- + H_3O^+$ und $B + H_2O \rightleftharpoons BH^+ + OH^-$, kommt man mit den Aktivitäten bzw. Konzentrationen der Teilchen zu den *Säure-* bzw. *Base-Dissoziationskonstanten*:

$K_S = a_{H_3O^+} \cdot a_{A^-}/a_{HA}$ und

$K_B = a_{BH^+} \cdot a_{OH^-}/a_B$ bzw.

$K^c_S = [H_3O^+] \cdot [A^-]/[HA]$ und

$K^c_B = [BH^+] \cdot [OH^-]/[B]$.

Für ein korrespondierendes Säure-Base-Paar gilt $K_S \cdot K_B = K_w$, so daß sich die Bestimmung einer der drei Größen erübrigt. Einige Säure- und Basekonstanten sind in Tab. 7.3.2-2 aufgeführt. Ihrer Definition nach handelt es sich um Dissoziationskonstanten. Gelegentlich werden auch Assoziationskonstanten tabelliert, z. B. für den Vergleich mit Komplexstabilitätskonstanten (↑ 7.3.2.6.); sie drücken die Assoziation von Brönstedbase und Proton bzw. die Stabilität des aus dem deprotonierten Teilchen und dem Proton gebildeten Komplexes aus: $K^c_{S(ass)} = [HA]/[H_3O^+] \cdot [A^-]$. Für einwertige

Tab. 7.3.2-2 pK_S- und pK_B-Werte einiger korrespondierender Säure-Base-Paare (für 22 °C)

pK_S	Säure	Base	pK_B
≈ −10	$HClO_4$	ClO_4^-	≈ +24
≈ −6	HCl	Cl^-	≈ +20
≈ −3	H_2SO_4	HSO_4^-	≈ +17
−1,75	H_3O^+	H_2O	+15,75
+1,9₀	HSO_4^-	SO_4^{2-}	+12,1₀
+1,9₅	H_3PO_4	$H_2PO_4^-$	+12,0₅
+1,9₅	"H_2SO_3" (SO_2/H_2O)	HSO_3^-	+12,0₅
+4,75	CH_3COOH	CH_3COO^-	+9,25
+6,5₀	"H_2CO_3" (CO_2/H_2O)	HCO_3^-	+7,5₀
+6,9₀	H_2S	HS^-	+7,1₀
+7,0	HSO_3^-	SO_3^{2-}	+7,0
+7,1₀	$H_2PO_4^-$	HPO_4^{2-}	+6,9₀
+9,25	NH_4^+	NH_3	+4,75
+9,4₀	HCN	CN^-	+4,6₀
+10,4	HCO_3^-	CO_3^{2-}	+3,6
+12,3	HPO_4^{2-}	PO_4^{3-}	+1,7
+12,9	HS^-	S^{2-}	+1,1
+15,74	H_2O	OH^-	−1,74
≈ +23	NH_3	NH_2^-	≈ −9

Säuren HA sind sie einfach die Kehrwerte der Dissoziationskonstanten, $pK = \lg K_{(ass)}$, bei mehrwertigen Säuren H_nA ist die reziproke Numerierung der einzelnen Protonierungsstufen zu beachten. So beschreibt die 1. Säuredissoziationskonstante $K_{S(1)}$ der Orthophosphorsäure H_3PO_4 deren Gleichgewicht mit $H_2PO_4^-$, die 1. Assoziationskonstante dagegen den Zusammentritt von Orthophosphat-Ionen PO_4^{3-} und Protonen zu HPO_4^{2-}.

In wäßriger Lösung sind stärkere Brönstedsäuren als das Hydroniumion, etwa Perchlor- und Schwefelsäure, und stärkere Brönstedbasen als OH^-, z. B. das Amidion NH_2^-, nicht existent, da weitgehende Umsetzung gemäß $HClO_4 + H_2O \rightarrow H_3O^+ + ClO_4^-$ bzw. $NH_2^- + H_2O \rightarrow NH_3 + OH^-$ stattfindet. Diese *Nivellierung* erklärt die Konstanz der auf unendlich verdünnte Lösungen bezogenen Neutralisationsenthalpie $\Delta_R H_\infty = -57{,}3$ kJmol^{-1} für die Umsetzung starker Säuren mit starken Basen in Wasser; sie kann in geeigneten anderen Lösungmitteln aufgehoben werden (↑ 7.3.2.5.).

Können Stoffe wie H_2SO_4, H_3PO_4 und H_2S mehrere Protonen abgeben, ist eine gesetzmäßige Verringerung deren Acidität mit abnehmender Zahl der noch abspaltbaren Protonen zu verzeichnen (↑ Tab. 7.3.2-2); entsprechend wächst die Basizität der konjugierten Brönstedbasen. Hierfür lassen sich strukturell-energetische und statistische Gründe angeben.

Jede Abspaltung eines Protons führt zu einer relativen Erhöhung der negativen Ladung der konjugierten Base; die Abspaltung immer weiterer Protonen wird durch die anwachsende Coulombanziehung erschwert, die überwunden werden muß. Entsprechend erleichtert wird die Protonenanlagerung an die jeweilige Base. Somit gilt: $K_{S(1)} > K_{S(2)} > \ldots > K_{S(n)}$. Für eine dreiwertige Säure H_3A ist die Abspaltung eines von drei Protonen wahrscheinlicher als die Anlagerung eines an den einen dafür verfügbaren Platz. Umgekehrt steht der Abspaltung des letzten von drei Protonen die Möglichkeit der Anlagerung eines Protons an einen von drei vorhandenen Plätzen an die konjugierte Base gegenüber. Auch hier wird $K_{S(1)} > K_{S(2)} > K_{S(3)}$. Bei mehrwertigen Säuren, deren acide Gruppen aber soweit voneinander entfernt sind, daß eine Coulombwechselwirkung weitgehend ausgeschlossen ist, werden vor allem die statistischen Aspekte dominieren, wie die höheren Dicarbonsäuren in Tab. 7.3.2-3 zeigen. Bei Molekülen wie H_3PO_4, wo alle aciden Gruppen an demselben Zentralatom gebunden sind, oder bei Molekülkonfigurationen organischer Säuren, die eine starke Wechselwirkung der aciden Gruppen ermöglichen, überwiegen die energetischen Gesichtspunkte. Von den Isomeren Fumar- und Maleinsäure stellt Fumarsäure die trans-Form dar (s. a. 5.2.2.2.). Die Carboxylgruppen sind hier hinreichend weit voneinander entfernt, ihre Dissoziation kann weitgehend unabhängig voneinander erfolgen. Die cis-Form Maleinsäure weist dagegen eine starke Differenzierung der pK_S-Werte auf (↑ Tab. 7.3.2-3). Ihre Acidität ist in erster Stufe erhöht, weil das dabei gebildete Anion durch die Ausbildung einer starken Wasserstoffbrücke stabilisiert wird, in zweiter Stufe erniedrigt, weil die Abspaltung des zweiten Protons aus der Brücke deutlich erschwert ist. Analog ergibt der Vergleich der pK_S-Werte der drei isomeren Benzendicarbonsäuren, daß die Phthalsäure das ortho-Isomere darstellen muß.

Die in Tab. 7.3.2-2 tabellierten Konstanten gestatten auch, durch Kombination der Werte für verschiedene Säure-Base-Paare die zwischen ihnen in wäßrigen Lösungen ablaufenden Reaktionen einzuschätzen. Aus den Teilschritten (1), $HA + H_2O \rightleftharpoons H_3O^+ + A^-$, und (2) $B + H_3O^+ \rightleftharpoons HB^+ + H_2O$, ergibt sich die Gesamtreaktion $HA + B \rightleftharpoons A^- + HB^+$; die Rolle des Lösungsmittels als Protonenüberträger ist hier gar nicht mehr ersichtlich. Entsprechend folgt aus den Gleichgewichtskonstanten der Teil-

Tab. 7.3.2-3 pK_S-Werte mehrwertiger Carbonsäuren bei 25 °C in wäßriger Lösung

	$pK_{S(1)}$	$pK_{S(2)}$	ΔpK
Aliphatische Dicarbonsäuren			
Oxalsäure $(COOH)_2$	1,27	4,27	3,0
Malonsäure $HOOC{-}CH_2{-}COOH$	2,86	5,70	2,84
Bernsteinsäure $HOOC{-}(CH_2)_2{-}COOH$	4,21	5,64	1,43
Glutarsäure $HOOC{-}(CH_2)_3{-}COOH$	4,34	5,27	0,93
Maleinsäure cis-Form $HOOC{-}CH{=}CH{-}COOH$	1,92	6,23	4,31
Fumarsäure trans-Form	3,02	4,38	1,36
Benzendicarbonsäuren $HOOC{-}C_6H_4{-}COOH$			
Phthalsäure (ortho-Form)	2,95	5,41	2,46
Isophthalsäure (meta-Form)	3,62	4,60	0,98
Terephthalsäure (para-Form)	3,54	4,46	0,92

reaktionen, $K_{S(1)} = a_{H_3O^+} \cdot a_{A^-}/a_{HA}$ und $K_{S(2)} = a_{H_3O^+} \cdot a_B/a_{HB^+}$, die Bruttogleichgewichtskonstante der Säure-Base-Reaktion
$K = a_{HB^+} \cdot a_{A^-}/(a_B \cdot a_{HA}) = K_{S(1)}/K_{S(2)}$;
$pK = pK_{S(1)} - pK_{S(2)}$.
Die Tabellenwerte ergeben, daß das Gleichgewicht einer Säure-Base-Reaktion um so weiter zugunsten der Reaktionsprodukte verschoben ist, je stärker sauer bzw. basisch die Ausgangsstoffe sind. Starke Säuren wie H_3O^+ (aus HCl) reagieren mit starken Basen wie OH^- (aus KOH) gemäß $pK_{(H_3O^+)} - pK_{(H_2O)} \approx -17{,}5$ vollständig unter Bildung von Wasser. Bei äquimolarem Vermischen von Säure und Lauge resultiert dann eine neutrale Lösung des entsprechenden Salzes, hier von Kaliumchlorid. Starke Säuren wie HCl reagieren mit mittelstarken bis schwachen Basen, etwa NH_3, mit $pK \approx -11$ ausreichend vollständig für eine analytische Anwendung der Reaktion; die bei 1 : 1-Umsatz entstehende Lösung reagiert aber schwach sauer. Die zur schwachen Base korrespondierende Säure, hier NH_4^+, neigt in weit stärkerem Maße zur Protonenabspaltung als die ausgesprochen schwache Anionbase Cl^- zur Anlagerung dieser Protonen. Entsprechend führt die Umsetzung starker Basen wie OH^- mit schwachen Säuren wie Essigsäure zu alkalischen Lösungen; das Acetation neigt stärker zur Protonenanlagerung als das zu OH^- konjugierte Wasser zur Protonenabspaltung.
Die Umsetzungen schwacher Säuren mit schwachen Basen erfolgen längst nicht so weitgehend wie in den vorgenannten Fällen. Die resultierende Lösung ist je nach den relativen Säure-Base-Stärken der Partner sauer oder alkalisch. Eine Ammoniumacetatlösung reagiert jedoch neutral, da zufällig beide Ionen die gleiche Tendenz haben, in die korrespondierenden Teilchen überzugehen. Der pK_S-Wert von Essigsäure stimmt mit pK_B von Ammoniak überein, dementsprechend auch pK_B der Anionbase Acetat mit pK_S der Kation-Säure NH_4^+ (↑ Tab. 7.3.2-2).
Insgesamt zeigen die Beispiele, daß der Ausdruck „Neutralisation" für solche Säure-Base-Reaktionen überholt ist; auch die Bezeichnung „Hydrolyse" für die Reaktion von Ionen eines durch Säure-Base-Reaktion entstandenen Salzes mit Wasser verliert bei Verwendung der Brönstedkonzeption die Sonderstellung, die ihr im Rahmen der Säure-Base-Definition nach ARRHENIUS als Umkehrreaktion der Neutralisation noch zukam.

7.3.2.4. Ermittlung des *p*H-Werts wäßriger Lösungen. In der Literatur werden zahlreiche Näherungsgleichungen für *p*H-Berechnungen angegeben, je nach den Aciditäts- und Basizitätsabstufungen bei den Reaktionspartnern. Dessenungeachtet kann man in allen Fällen einheitlich vorgehen, um für die auftretenden Unbekannten hinreichend viele unabhängige Bestimmungsgleichungen zu erhalten. Bei einwertigen Säuren oder Basen lassen sich vier, bei mehrwertigen entsprechend der Zahl der zusätzlichen Dissoziationsschritte ebenso viele weitere Gleichungen angeben. Für die Umsetzung einer zweiwertigen Säure H_2A mit einer Alkalilauge MOH zu MHA und M_2A ergeben sich z. B. (c_{H_2A} Totalkonzentration an H_2A; ϱ Reaktionsgrad = 0 bis 2; $\varrho \cdot c_{H_2A} = [M^+]$):

Massenbilanz: $c_{H_2A} = [H_2A] + [HA^-] + [A^{2-}]$

Ladungsbilanz (Elektroneutralitätsbeziehung): $[H^+] + \varrho \cdot c_{H_2A} = 2[A^{2-}] + [HA^-] + [OH^-]$,

Säuredissoziationskonstanten $K_{S(1)}$ und $K_{S(2)}$:
$K_{S(1)} = [HA^-] \cdot [H^+]/[H_2A]$;
$K_{S(2)} = [A^{2-}]\,[H^+]/[HA^-]$,

Autoprotolysekonstante des jeweiligen Lösungsmittels: $K_w^c = [H^+] \cdot [OH^-]$.
Diese fünf Gleichungen reichen zur Bestimmung der unbekannten Größen $[H^+]$, $[H_2A]$, $[HA^-]$, $[A^{2-}]$ und $[OH^-]$ aus; c_{H_2A} wird vorgegeben, der Reaktionsgrad ϱ ist durch den jeweiligen Laugenverbrauch bekannt, auch die Werte K_w, $K_{S(1)}$ und $K_{S(2)}$ liegen im allgemeinen vor. Mißt man den *p*H-Wert in Abhängigkeit vom Umsetzungsgrad ϱ bei bekanntem c_{H_2A}, lassen sich natürlich auch die Konstanten $K_{S(1)}$ und $K_{S(2)}$ ermitteln. Bei nicht zu kleinen Totalkonzentrationen c_{H_2A} vereinfachen sich die Rechnungen für mittlere *p*H-Bereiche dadurch, daß in Summen oder Differenzen auftretende Konzentrationen $[H^+]$ und $[OH^-]$ vernachlässigt werden können. Liegen die pK_S-Werte hinreichend weit auseinander (z. B. bei der schon behandelten Maleinsäure) können die beiden Dissoziationsschritte getrennt behandelt werden. Dadurch wird die Rechnung auf den einfachen Fall der nacheinander ablaufenden Dissoziation zweier einwertiger Säuren HA reduziert. Die Verhältnisse lassen sich anschaulich mit Hilfe der von BJERRUM (1940) eingeführten *Komplexbildungsfunktion* $\bar{n}$ darstellen,

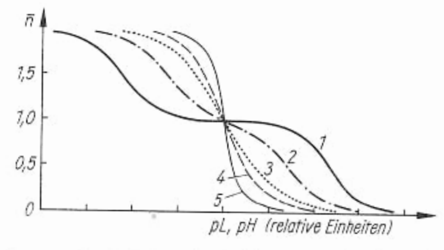


Abb. 7.3.2-4 Komplexbildungsfunktion $\bar{n}$ nach BJERRUM für Säuren H_2A bzw. Metallkomplexe ML_2 für unterschiedliche Quotienten der Gleichgewichtskonstanten K_1 und K_2; K_1/K_2 beträgt 40000 für *1*, 400 für *2*, 16 für *3*, 4 für *4* und 0 für *5*

die die durchschnittliche Anzahl der pro Zentralteilchen gebundenen Liganden [*ligare* lat., binden] angibt. Es lassen sich damit beliebige Stufengleichgewichte beschreiben, z. B. die Protonierung von Brönstedbasen A oder die Bindung von n Liganden L an Metallionen M. Weder die Ladungen der korrespondierenden Teilchen noch die Zahl n, die auch größere Werte als 2 annehmen kann, sind von Belang. Für den Rechenaufwand spielt es allerdings eine große Rolle, ob die einzelnen Schritte deutlich getrennt (Abb. 7.3.2-4, Kurven *1* und *2*) oder gekoppelt (Kurven *3* und *4*) ablaufen. Bei hinreichender Trennung können die entsprechenden Werte $pK_{S(2)}$ und $pK_{S(1)}$ mit den pH-Werten für Protonierungsgrade $\bar{n} = 0{,}5$ und $1{,}5$ gleichgesetzt werden. Die genannte Maleinsäure entspricht etwa Kurve *1* der Abbildung, die Fumarsäure Kurve *3*.

Sind die Konstanten für einzelne Assoziationsstufen eng benachbart (vgl. Kurve *4* in Abb. 7.3.2-4), was z. B. bei mehrwertigen Aminocarbonsäuren häufig auftritt, lassen sich zwar ebenfalls rechnerisch konsistente Säurekonstanten $K_{S(1)} \dots K_{S(n)}$ ermitteln, jedoch ist die Zuordnung der einzelnen Konstanten zu bestimmten aciden Gruppen selten zweifelsfrei möglich. Sofern nicht spektroskopische oder andere unabhängige Methoden weiterhelfen, kann man versuchen, durch chemische Blockierung einzelner acider Gruppen, etwa durch deren Veresterung oder Alkylierung, die Protonenzahl im Molekül zu verringern, ohne die Substituentenwirkung auf die verbleibenden aciden Funktionen allzu stark zu verändern. Der Vergleich der dann erhaltenen K_S-Werte mit denen der Stammsubstanz ermöglicht eine bessere Einschätzung des Struktureinflusses auf die Reaktivität bei Säure-Base- und damit zusammenhängenden Reaktionen, z. B. bei biokatalytischen Reaktionen.

Experimentell zugänglich sind pH-Werte von Lösungen durch zahlreiche Methoden, am direktesten durch elektrochemische Messungen mittels Wasserstoff- und entsprechender Bezugselektrode; praktisch einfacher als der Einsatz der *Wasserstoffelektrode* gestaltet sich die Verwendung einer auf die Protonenaktivität ansprechenden Glaselektrode (s. a. 7.3.4.).

Für die nur ungefähre Ermittlung des vorliegenden pH-Werts oder für die Anzeige größerer pH-Änderungen reichen oft geringe Zusätze von *pH-Indikatoren* zur Lösung aus. Hierbei handelt es sich meist um organische Farbstoffe HI geringe Acidität, deren korrespondierende Base I eine andere Farbe besitzt. Sie werden so ausgewählt, daß ihr pK_S-Wert dem pH-Wert im Äquivalenzpunkt der zu verfolgenden Säure-Base-Reaktion entspricht. Im allgemeinen liegen dann bei $pH = pK_S$ gerade gleiche Konzentrationen der unterschiedlich gefärbten Teilchen HI und I des Indikators vor. Bei einer geringfügigen Über- oder Unterschreitung des Äquivalenzpunkts wird dieses Verhältnis zugunsten der Farbform I bzw. HI stark verschoben, was am Farbumschlag zu erkennen ist.

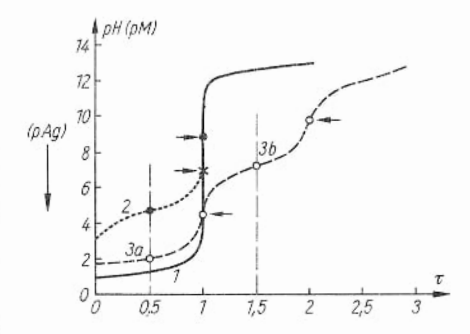


Abb. 7.3.2-5 Titrationskurven für ausgewählte Säuren; pH-Wert in Abhängigkeit von der Zahl τ der Äquivalente KOH, Ausgangskonzentration 0,1 mol l^{-1}, bei Raumtemperatur; *1* HCl (Bromthymolblau 6,2–7,6), *2* Essigsäure (Phenolphthalein 8,0–10,0), *3* H_3PO_4 (*a* Bromkresolgrün 4,0–5,6, *b* Thymolphthalein 9,4 bis 10,6); bzgl. pM, pAg s. a. 7.3.2.6.

Abb. 7.3.2-5 zeigt einige Umsetzungskurven ein- und mehrwertiger Säuren bei Raumtemperatur, wie sie durch potentiometrische pH-Titration mittels der sehr starken Arrheniusbase KOH, d. h. der Brönstedbase OH^-, erhalten werden oder sich mit den zuvor angegebenen Gleichungen berechnen lassen. Geeignete pH-Indikatoren zur Anzeige der durch Pfeile markierten Äquivalenzpunkte sind ebenfalls angegeben. Bei stärker konzentrierten Lösungen würden die pH-Sprünge noch ausgeprägter, bei wesentlich verdünnteren, etwa 10^{-4} molaren Lösungen, ergäben sich lediglich bei den stärksten Säuren noch zuverlässig auswertbare Kurven.

An den Kurven *2* und *3b* der Abb. 7.3.2-5 ist zu erkennen, daß bei mittleren Umsetzungsgraden um 0,5 und 1,5, wo etwa gleiche Anteile der entsprechenden Säure-Base-Paare vorliegen, die pH-Änderungen bei Laugenzusatz am geringsten sind. Solche etwa äquimolaren Mischungen von Säure und korrespondierender Base reagieren auf Zugabe von Protonen oder Hydroxidionen relativ wenig, sie kompensieren sie durch Gleichgewichtsverschiebung weitgehend. Solche *Pufferlösungen* haben also ihre maximale Pufferkapazität jeweils bei $pH = pK_S$. Aus Tab. 7.3.2-2 lassen sich weitere geeignete Puffersysteme entnehmen, z. B. für pH-Werte zwischen 9 und 10 Lösungen von Ammoniak/Ammoniumchlorid. Die pH-Pufferung ist be-

sonders für biochemische Vorgänge bedeutsam; z. B. wird durch den Puffer $CO_2(H_2O)/HCO_3^-$ der *p*H-Wert des Bluts auf etwa 7,4 eingestellt.

7.3.2.5. Abhängigkeit der Säurestärke vom Lösungsmittel. Für die Acidität eines bestimmten Stoffs HA ist neben der Basizität des Lösungsmittels auch dessen Dielektrizitätskonstante ε wichtig, da die Gleichgewichte je nach der Ladung der korrespondierenden Teilchen unterschiedlich stark auf ε ansprechen. So findet bei der Dissoziation von Kationsäuren, etwa gemäß $NH_4^+ + H_2O \rightleftharpoons NH_3 + H_3O^+$, lediglich eine Übertragung der positiven Ladung von einem auf ein anderes Teilchen statt. Hier ist ε von viel geringerem Einfluß als im Falle der Dissoziation einer Neutralsäure wie Essigsäure, die gemäß $Hac + H_2O \rightleftharpoons ac^- + H_3O^+$ das Entstehen und die Trennung entgegengesetzter Ladungen einschließt. Eine hohe Dielektrizitätskonstante wird eine solche Dissoziation begünstigen. Tab. 7.3.2-6 gibt einige Beispiele für Lösungsmittel an, die trotz recht unterschiedlicher Eigenschaften als wasserähnlich anzusprechen sind, weil sie ebenfalls eine Autoprotolyse erfahren.

Neben der Autoprotolyse- und Dielektrizitätskonstanten ist der saure oder basische Charakter der Lösungsmittel zu beachten, um die Relativverschiebung der Säure-Base-Gleichgewichte gegenüber denen in wäßriger Lösung abschätzen zu können. Eisessig, also wasserfreie Essigsäure, ist z. B. stärker sauer als Wasser, zeigt aber nur eine geringe Neigung zur Protonenaufnahme. Es kommt folglich zu einer differenzierten Reaktion auch mit Säuren wie $HClO_4$ oder HCl, die in Wasser wegen dessen stärkerer Basizität nivelliert werden (↑ 7.3.2.3.). Statt dessen wird in Eisessig ein weitgehender Ausgleich der Basenstärke basischer Stoffe erreicht, da das relativ stark acide Molekül CH_3COOH auch an sehr schwache Basen noch Protonen überträgt. Deshalb lassen sich in Eisessig noch Basen quantitativ bestimmen, z. B. durch Titration mit Perchlorsäure, für die in wäßriger Lösung eine Indizierung des Äquivalenzpunkts wegen der nur äußerst unvollständig verlaufenden Reaktion gar nicht möglich wäre.

Flüssiges Ammoniak ist ein stark basisches Lösungsmittel, führt also zur Nivellierung der Säurestärke auch bei weniger starken Säuren auf die Acidität des Ammoniumions NH_4^+.

Alkohole unterscheiden sich in Acidität und Basizität relativ wenig von Wasser. Sie werden oft – allein oder im Gemisch mit Wasser – eingesetzt, wenn in reinem Wasser die zu untersuchenden Stoffe zu wenig löslich sind.

Werden mehrere Säure-Base-Paare von einheitlichem Ladungstyp miteinander verglichen, bleiben ihre Relativpositionen für verschiedene Lösungsmittel und damit für verschiedene Aciditätsskalen weitgehend erhalten. Das gilt aber wegen des Einflusses von ε nicht mehr, wenn für unterschiedliche Ladungstypen eine z. B. in Wasser erhaltene Rangfolge mit der in einem anderen Medium verglichen wird.

Eine solche Modifizierung geht z. B. aus Abb. 7.3.2-7 hervor, wo eine schwache organische

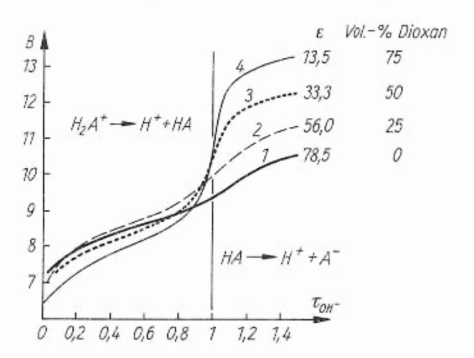


Abb. 7.3.2-7 Lösungsmitteleinfluß auf die Dissoziation eines Aminoximhydrochlorids H_2A^+ in Wasser-Dioxan-Mischungen; *B* *p*H-Meter-Ablesung, τ_{OH^-} Umsetzungsgrad mit KOH

Tab. 7.3.2-6 Wasser und einige wasserähnliche Lösungsmittel; ϑ_F, ϑ_V Schmelz- und Siedetemperatur in °C, ε Dielektrizitätskonstante, σ spezifische Leitfähigkeit in Ω^{-1} cm^{-1}, K_{HY} Autoprotolysekonstante in mol² l^{-2}; in Klammern die jeweiligen Meßtemperaturen

Flüssigkeit	ϑ_F	ϑ_V	ε		σ		K_{HY}	
Wasser	0	100	78,5	(25)	$4,4 \cdot 10^{-8}$	(18)	10^{-14}	(22)
Methanol CH_3OH	−97,5	64,7	32	(18)	$4,4 \cdot 10^{-7}$	(18)	10^{-17}	(25)
Ethanol C_2H_5OH	−114,5	78,3	26	(18)	$6,4 \cdot 10^{-8}$	(18)	10^{-19}	(25)
Essigsäure	16,7	118,5	6,2	(25)	$5 \cdot 10^{-9}$	(25)	10^{-13}	(25)
Ammoniak	−77,7	−33,4	22	(−34)	$1,3 \cdot 10^{-7}$	(−79)	10^{-22}	(−33)
Blausäure HCN	−14,7	26	158	(0)	$3,3 \cdot 10^{-6}$	(0)		

Säure H_2A^+ in Wasser bzw. in Wasser-Dioxan-Mischungen unterschiedlichen Dioxangehalts mit Kalilauge titriert wurde. Der cyclische Ether Dioxan weist wie andere Ether eine gewisse Tendenz zur Protonenanlagerung auf, ohne unter den üblichen Bedingungen selbst Protonen abspalten zu können. Neben der dadurch und durch Änderungen der Solvatationsneigung bewirkten Modifizierung der Lösungsmitteleigenschaften wirkt sich vor allem die Änderung der Dielektrizitätskonstante auf die Säure-Base-Gleichgewichte aus. Während sich die Dissoziation als Kationsäure wenig ändert, wird die Abspaltung des zweiten Protons mit wachsendem Dioxangehalt zunehmend behindert. Dadurch kommt es zu einer deutlichen Differenzierung von $K_{S(1)}$ und $K_{S(2)}$, d. h. zu einer Trennung der beiden Dissoziationsschritte.

7.3.2.6. Komplex-, Fällungs- und Redoxgleichgewichte in wäßrigen Lösungen. Die für Säure-Base-Gleichgewichte geltenden Zusammenhänge können auch für das Verständnis von Komplex-, Fällungs- und Redoxgleichgewichten herangezogen werden, obwohl es für die Abhängigkeit von Komplexstabilität, Löslichkeit und Redoxstabilität von der Struktur der Stoffe nur gelegentliche Gemeinsamkeiten mit den Säure-Base-Gleichgewichten gibt. Die deutlichsten Bezüge zu Säure-Base-Reaktionen haben die Komplexgleichgewichte.

7.3.2.6.1. *Komplexgleichgewichte.* Wie bereits in 7.3.2.3. erwähnt, können auch Protonenaustauschreaktionen als Bildung oder Zerfall von Wasserstoffkomplexen beschrieben werden. Bei der Bildung von Metallkomplexen in wäßriger Lösung handelt es sich ebenfalls um einen Austauschvorgang. Die Wassermoleküle in der Hydrathülle der Metallionen werden durch die stärkeren Liganden ebenso ersetzt wie stärker basische Teilchen ein Proton von einem schwächer basischen übernehmen können.

Für Stabilitätskonstanten von Komplexen MA_n, $K_1^c = [MA]/([M]\,[A])$; $K_2^c = [MA_2]/([MA]\,[A])$; ... $K_n^c = [MA_n]/([MA_{n-1}]\,[A])$, und *Bruttostabilitätskonstante*, $\beta_n^c = [MA_n]/([M]\,[A]^n)$ $= K_1^c \cdot K_2^c \ldots \cdot K_n^c$, ergibt sich häufig eine Proportionalität zu entsprechenden Säureassoziationskonstanten der Bildung protonierter Teilchen mit demselben Liganden, z. B. von HA^{2-}, H_2A^-, H_3A.

Abb. 7.3.2-8 Diacetyldioxim (DH_2) und die Struktur des Komplexes $Ni(DH)_2$

Neben statistischen Faktoren wirken auch spezifisch strukturell-energetische auf die Komplexstabilität. So wird der quadratisch-planare 1 : 2-Komplex $Ni(DH)_2$ des analytisch als Nickelreagens bedeutsamen Diacetyldioxims (Abb. 7.3.2-8) durch die Ausbildung starker Wasserstoffbrückenbindungen erheblich stabilisiert gegenüber dem 1 : 1-Komplex $Ni(DH)^+$. Jedes Molekül des Liganden DH^- betätigt zwei Haftstellen gegenüber dem Nickel(II)-ion, wirkt also als zweizähliger oder auch zweizähniger Ligand.

Abb. 7.3.2-9 Dinatriumsalz der EDTE, Na_2H_2Y, und ein oktaedrisch konfigurierter Metallkomplex MY (schematisch)

Besitzen Komplexliganden eine größere Zahl von sterisch günstig zueinander orientierten Haftstellen, werden durch deren gleichzeitige Fixierung an das Zentralatom M besonders stabile Komplexe gebildet. Es entfällt dadurch zugleich die bei einzähligen Liganden wie NH_3 oder den Halogenidionen oft störende unzureichende Abgrenzung der einzelnen Komplexbildungsschritte voneinander. Beispielsweise tragen Moleküle der durch Schwarzenbach untersuchten Ethylendiamintetraessigsäure (EDTE) der allgemeinen Formel H_4Y in ihrer deprotonierten Form Y^{4-} sechs Haftstellen, die im allgemeinen gerade für die Bildung eines 1 : 1-Komplexes günstig sind (Abb. 7.3.2-9). Unter geeigneten *p*H-Bedingungen bilden sich in einem Schritt farblose Metall-EDTE-Komplexe wie ZnY^{2-}, CaY^{2-} usw. Die Reaktion $M^{n+} + H_2Y^{2-} \rightleftharpoons MY^{(4-n)-} + 2\,H^+$ wird in der *Komplexometrie* zur Bestimmung des Metallgehalts von Lösungen ausgenutzt. Das Prinzip läßt sich an Abb. 7.3.2-5 erläutern, wenn man deren Achsenbezeichnung *p*H durch *p*M ersetzt. Anfangs liegen solvatisierte Metallionen, vereinfacht als M bezeichnet, in beträchtlicher Konzentration [M] vor, der entsprechende *p*M-Wert ist also niedrig. Durch Zugabe des Komplexbildners EDTE, meist in der Form H_2Y^{2-},

werden die Metallionen gebunden und die dabei freigesetzten Protonen durch pH-Puffer entfernt, so daß das Gleichgewicht in der gewünschten Richtung verschoben wird. Dadurch erhöht sich der pM-Wert, zunächst nur allmählich, bis bei Annäherung an den Äquivalenzpunkt durch die weitgehende Bindung der Metallionen ein deutlicher pM-Sprung einsetzt. Er kann durch *Metallindikatoren* indiziert werden, die in ihrer Funktion das Gegenstück zu den pH-Indikatoren darstellen. Es sind Farbstoffe, die ebenfalls Komplexe mit den betreffenden Metallen bilden, wobei sich die Farbe des Komplexes deutlich von der des freien Farbstoffs unterscheiden soll. Zur richtigen Endpunktsanzeige muß der Metall-Indikator-Komplex MI weniger stabil sein als der von EDTE selbst gebildete, aber doch wesentlich stärker als der Aquokomplex $[M(OH_2)_x]^{n+}$, damit aus ihm erst gegen Ende der Reaktion, also nach Aufzehren des Aquokomplexes, der Indikator unter Farbumschlag frei wird. Wie bei den pH-Titrationen geht es hier darum, den Zusammenhang zwischen Ligandstruktur und Komplexstabilität zur Auswahl optimaler Indikatoren zu untersuchen.

7.3.2.6.2. *Fällungsreaktionen.* Wegen der sehr komplexen Zusammenhänge zwischen Zusammensetzung, Struktur und Löslichkeit von Stoffen ist es kaum möglich, Löslichkeitsabstufungen für Substanzreihen in Wasser oder anderen Lösungsmitteln sicher vorauszusagen. Gittertyp, Ionensolvatation, Ionenladung und -größe, Bindungspolarität und Teilchenpolarisierbarkeit sind einige der Eigenschaften, die hier zusammenwirken.

Quantitative Vergleiche zur Löslichkeit ionischer Verbindungen AB gestattet das *Löslichkeitsprodukt* (↑ 7.1.1.4.) $L_{AB} = a_{A^+} \cdot a_{B^-}$ bzw. $L^c_{AB} = [A^+]\,[B^-]$, das bei Stoffen anderer Stöchiometrie A_xB_y in der allgemeineren Form $L^c_{A_xB_y} = [A]^x \cdot [B]^y$ zu schreiben ist. Bei identischem Formeltyp sind die Löslichkeitsprodukte unmittelbar vergleichbar, bei unterschiedlicher Stöchiometrie sind die Gleichgewichtskonzentrationen für die jeweils interessierenden Teilchen, z. B. für A^{n+}, gegenüberzustellen. Bevorzugt wird A^{n+} in *den* Niederschlag übergehen, zu dem unter den gegebenen Bedingungen die geringste Gleichgewichtskonzentration $[A^{n+}]$ gehört. Über diese Größe können auch Kopplungen mit Säure-Base- oder Komplexgleichgewichten quantitativ erfaßt werden.

Ist das Teilchen A z. B. Partner eines günstig gelagerten Säure-Base-Gleichgewichts, kann es aus dem Löslichkeitsgleichgewicht entfernt, durch die Gleichgewichtsverschiebung der Niederschlag also aufgelöst werden. So treten die über einem Bariumcarbonatbodenkörper in Lösung vorhandenen Carbonationen mit den Protonen einer zugesetzten Mineralsäure über die Stufe des Hydrogencarbonats HCO_3^- zu Wasser und gasförmig entweichendem CO_2 zusammen, bis die gesamte Fällung aufgelöst ist. Das schwerlösliche Silberchlorid AgCl mit $L \approx 10^{-10}$ mol²l⁻² wird durch wäßriges Ammoniak deshalb gelöst, weil ein recht stabiler Silber-diammin-Komplex $[Ag(NH_3)_2]^+$ mit $\lg \beta^c_2 = 7$ gebildet werden kann, der mit einer geringeren Konzentration $[Ag^+]$ im Gleichgewicht steht als in Wasser suspendiertes AgCl. Bei AgI mit seinem gegenüber AgCl um etwa sechs Zehnerpotenzen kleineren Löslichkeitsprodukt stehen dagegen nicht ausreichend viele Silberionen in Lösung zur Verfügung, um nennenswerte Mengen des Diamminkomplexes zu bilden. Die Auflösung auch dieser Fällung gelingt aber durch Behandeln mit Kaliumcyanidlösung, da die Stabilität des Dicyano-argentat-Komplexes $[Ag(CN)_2]^-$ ihrerseits um etwa 30 Größenordnungen über der des Diamminkomplexes liegt.

Die analytische Ausnutzung von Löslichkeitsgleichgewichten ist über *Fällungstitrationen* möglich, die ebenfalls an Abb. 7.3.2-5 erläutert werden können. Eine getrennte Bestimmung zweier Halogenide, z. B. von Chlorid und Iodid durch Fällung mit Silbernitratlösung, ähnelt der pH-Titration einer zweiwertigen oder der zwei ersten Stufen einer dreiwertigen Säure mit Lauge; lediglich die Achse für den Ionenexponenten pAg ist gegenläufig zu der für pH. Bei Zugabe der ersten Anteile $AgNO_3$ fällt das am schwersten lösliche AgI aus, die Konzentration $[Ag^+]$ ist dementsprechend sehr gering, der pAg-Wert relativ hoch. Bei weitgehender Ausfällung des Iodids tritt ein Sprung zu kleineren pAg-Werten ein, bis die zur Ausfällung von AgCl notwendige Konzentration $[Ag^+]$ erreicht ist. Mit erfolgter AgCl-Ausfällung sinkt pAg ein weiteres Mal sprunghaft, entsprechend dem sich ergebenden Überschuß an Fällungsmittel.

Die Indizierung des Endpunkts kann elektrochemisch erfolgen oder durch *Fällungsindikatoren*, die in Löslichkeit oder Komplexstabilität so auszuwählen sind, daß ihr Farbumschlag gerade bei den Ionenexponenten pAg oder pX (X = Cl, Br, I) der jeweiligen Äquivalenzpunkte erfolgt. Eine getrennte Bestimmung verschiedener Ionen gelingt aber – ebenso wie die gesonderte Erfassung von Säuredissoziations- oder Komplexbildungsstufen – nur bei hinreichenden Abstufungen, hier in der Gleichgewichtskonzentration der jeweiligen Teilchen. Ist das z. B. bei der Titration von Cl^-/Br^- wegen $L_{AgBr} \approx 10^{-12}$ mol²l⁻² nur unzureichend der Fall, könnten zwar rechnerisch Korrekturen zur Angleichung des erhaltenen an den wahren Äquivalenzpunkt vorgenommen werden, die Verhältnisse werden aber oft durch *Mitfällung*, d. h. partielle Mischkristallbildung, kompliziert.

Außerdem sind die variablen Objektmengenverhältnisse der interessierenden Stoffe zu berücksichtigen; die bei Protonierungs- oder anderen Stufengleichgewichten eines Stoffs zwangsweise vorliegende Äquimolarität der konjugierten Teilchen ist hier höchstens zufällig einmal gegeben.

7.3.2.6.3. *Redoxgleichgewichte.* Sie sind immer das Ergebnis der Umsetzung unterschiedlich starker Oxydations- oder Reduktionsmittel miteinander. Im Unterschied zu den anderen genannten Gleichgewichten, bei denen es sich um die Kombination von Ionen handelt, werden hier Elektronen ausgetauscht. Häufig werden Redoxgleichgewichte aber zugleich von Säure-Base-Gleichgewichten und Komplexbildungsreaktionen, gelegentlich auch von Fällungsreaktionen überlagert.

So wie in wäßrigen Lösungen keine freien Protonen auftreten, existieren dort auch keine freien Elektronen. In Analogie zu Protolysereaktionen sind deshalb auch bei Redoxreaktionen Elektronenaufnahme und -abgabe über *korrespondierende Redoxpaare* miteinander gekoppelt. Eine ausführliche Behandlung der Spezifik des Elektronenaustauschs und der sich – oft über Phasengrenzflächen hinweg – einstellenden Gleichgewichte erfolgt in Abschnitt 7.3.4.

Die Bestimmung redoxaktiver Teilchen kann durch *Redoxtitration* erfolgen. Entsprechend den Standardelektrodenpotentialen für die analytisch zu erfassenden Redoxpaare, etwa $U_H^{\ominus}(Fe^{2+}/Fe^{3+}) \approx 0{,}8$ V, wählt man ein Reagens mit hinreichend höherem $U_H^{\ominus}$-Wert, zum Beispiel eine Cerium(IV)-sulfatlösung mit $U_H^{\ominus}(Ce^{3+}/Ce^{4+}) \approx 1{,}6$ V, um eine praktisch vollständige Oxydation von Fe(II) zu Fe(III) nach $Ce^{4+} + Fe^{2+} \rightarrow Ce^{3+} + Fe^{3+}$ zu erzielen, auch hier wieder in formaler Entsprechung zu einer pH-Titrationskurve († Abb. 7.3.2-5, Fall *1*). Der *Redoxindikator* ist so zu wählen, daß seine Farbänderung durch Oxydation möglichst im Äquivalenzpunkt erfolgt, im gewählten Beispiel also etwa bei 1,2 V.

7.3.3. Reaktionen mit Feststoffen und an Grenzflächen

Reaktionen, an denen eine oder mehrere feste Phasen beteiligt sind, weisen wegen der relativ starren Fixierung der Teilchen im periodischen wie auch aperiodischen Gitter Besonderheiten auf gegenüber Umsetzungen zwischen Gasen oder miteinander mischbaren Flüssigkeiten. Im Unterschied zu Flüssigkeiten und Gasen liegen Feststoffe auch selten im an sich *normalen Zustand* des thermodynamischen Gleichgewichts vor, das als Defekte nur die *Eigenfehlordnung* zuläßt († 6.4.7.3.1.). Meist existieren zusätzliche Nichtgleichgewichtsdefekte, unterschiedlich in ihrer Natur († 6.4.7.) und Wirkung auf die Reaktivität der Feststoffe, die in ihrer Gesamtheit den normalen in einen *aktiven Zustand* überführen.

Das Verständnis von Feststoffreaktionen verlangt die Aufklärung einer ganzen Schrittfolge, wobei physikalische Vorgänge des Stoff-, Energie- und Impulstransports ebenso von Bedeutung sind wie die eigentliche chemische Reaktion. Als Teilschritte heterogener Reaktionen kommen vor allem in Betracht: Diffusion der Ausgangsstoffe aus dem Phaseninnern an die Grenzfläche – Adsorption des Reaktionspartners an der Oberfläche – chemische Reaktion im Bereich der Grenzfläche – Desorption der Endstoffe – Abdiffundieren der Reaktionsprodukte.

Der Stofftransport vollzieht sich als Platzwechsel einzelner Gitterbausteine, entweder durch die Gitter- oder *Volumendiffusion*, die *Korngrenzendiffusion* und die *Oberflächendiffusion* oder durch chemischen Transport über die Gasphase. Durch den jeweiligen Diffusionsmechanismus und die zugehörige Aktivierungsenergie des Platzwechselvorgangs wird die Reaktionsgeschwindigkeit bestimmt. Sie ist im allgemeinen erst bei relativ hohen Temperaturen nennenswert, so daß die Feststoffreaktionen – zumindest die, an denen ausschließlich feste Phasen teilnehmen – ihre Domäne bei thermisch stabilen anorganischen Verbindungen haben.

Das heißt nicht, daß Platzwechselvorgänge nicht auch bei relativ niedrigen Temperaturen nachweisbar wären. Von Tammann wurde die „Temperatur des merklich beginnenden Platzwechsels" definiert, bei der ein Zusammenbacken der Kristallite eines feinen Pulvers des jeweiligen Stoffs zu beobachten ist. Die *Tammanntemperaturen* liegen für Metalle bei etwa 1/3, für Oxide bei etwa 2/3, für organische Verbindungen oberhalb 3/4 der Schmelztemperatur T_F dieser Stoffe. Dieses Sinterverhalten kann durch chemische Fehlordnung († 6.4.7.2.) stark beeinflußt werden. Zusatz von etwas Lithiumoxid Li_2O zu Zinkoxid erhöht z. B. die Dichte von Sauerstoffehlstellen und erleichtert die Anionendiffusion, also den langsamsten und damit geschwindigkeitsbestimmenden Schritt des Sinterns. Entsprechend erniedrigt die Einführung von Oxiden dreiwertiger Metalle, wie Al_2O_3 und Ga_2O_3, durch Abbau von Sauerstoffehlstellen die Geschwindigkeit des Platzwechsels.

Gemäß 6.4.8.2. werden Platzwechselvorgänge bevorzugt durch *Leerstellendiffusion* realisiert; energetisch sind sie besonders an Korngrenzen und noch stärker an der Oberfläche von Kristallen begünstigt. Die Diffusionskoeffizienten sind bei Korngrenzendiffusion um etwa drei bis vier, bei Oberflächendiffusion um weitere drei Zehnerpotenzen größer als bei Gitterdiffusion. Da-

mit verliert der letztere, bei Einkristallen dominierende Vorgang in polykristallinem Material mit abnehmender Größe der Kristallite immer mehr an Bedeutung. Neben der starken Abhängigkeit der Festkörperreaktivität von der Realstruktur der Kristalle, ihrer Korngröße und Korngeometrie, ergibt sich daraus zwanglos der innere Zusammenhang zwischen Feststoffreaktionen und Umsetzungen an Oberflächen oder anders gearteten Grenzflächen.

7.3.3.1. Einteilung von Feststoffreaktionen. Feststoffreaktionen gehören zu den *heterogenen Reaktionen*. Das sind Umsetzungen, an denen mindestens zwei verschiedene Phasen beteiligt sind. Nach den Aggregatzuständen fest *s*, flüssig *l* und gasförmig *g*, in denen die Stoffe vorliegen, können folgende heterogene Reaktionen unterschieden werden: *g–l*, *l–l*, *g–s*, *l–s* und *s -s*. Die beiden ersten Fälle stehen den Reaktionen in homogener Phase nahe; auch bei ihnen liegen die Diffusionskoeffizienten der Teilchen meist um mehrere Größenordnungen über denen bei Feststoffreaktionen, also den heterogenen Reaktionen mit Beteiligung wenigstens *einer* festen Phase.

Für die *Reaktionen gasförmig–fest* ist die Struktur der Festkörperoberfläche für die auf ihr stattfindenden Ad- und Desorptionsprozesse von ausschlaggebender Bedeutung. Wichtige Sonderfälle sind die heterogen-katalytischen Reaktionen der chemischen Technik (↑ 7.3.3.3.) und die Oxydation von Metallen. Bilden sich bei letzterer dünne, am Auftreten von Interferenzfarben erkennbare Deckschichten, spricht man von *Anlaufvorgängen*, wachsen dicke Reaktionsschichten, von *Zundervorgängen*.

Reaktionen flüssig–fest laufen ab beim Auflösen von Salzen, bei der Reaktion von Metallen mit Säuren oder Laugen und bei der sogenannten *Flotation* zur Erzaufbereitung. Hier sind auch die elektrochemischen Prozesse zur Metallgewinnung bzw. -reinigung zu nennen, die wegen ihrer besonderen Bedeutung und ihrer durch den Transport elektrischer Ladung bedingten Spezifik aber gesondert behandelt werden (↑ 7.3.4.).

Reaktionen fest–fest werden häufig als Feststoffreaktionen im engeren Sinne bezeichnet. Die einfachsten Beispiele sind die allotropen bzw. polymorphen Umwandlungen von Elementen bzw. Verbindungen (s. a. 6.4.5.2.). Verwickelter verlaufen die *additiven Reaktionen*, bei denen aus einfachen Oxiden, Sulfiden, Halogeniden usw. komplexere Verbindungen wie Spinelle und Perowskite (↑ Tab. 6.4.4-7) entstehen. Gleiches gilt für die *doppelten Umsetzungen*, z. B. zwischen unterschiedlichen Halogeniden zweier Metalle. Der Vorgang der *Dotierung* zur Herstellung von Störhalbleitern oder von Luminophoren ist oft ebenfalls eine Feststoffreaktion im engeren Sinne.

7.3.3.2. Mechanismen ausgewählter Feststoffreaktionen. Der Ablauf von Feststoffreaktionen läßt sich auf die in der Reaktionskinetik (↑ 7.2.) übliche Weise durch die Ermittlung des Reaktionsgrads ϱ in Abhängigkeit von der Zeit charakterisieren. Zur Ermittlung der jeweiligen Zusammensetzung des Systems werden neben den üblichen chemischen Methoden der Konzentrations- und Spurenanalytik bevorzugt röntgendiagnostische (↑ 6.4.2.3.) und thermoanalytische Verfahren eingesetzt. Daneben können z. B. auch Messungen der Dichte, des thermischen Ausdehnungskoeffizienten, des Magnetismus und der elektrischen Leitfähigkeit des Gemischs Aufschluß über das Fortschreiten der Reaktion geben.

Zwar gestattet die mathematische Formulierung der Umsatzkurven – wie generell in der Kinetik – keine eindeutigen Schlüsse auf die eigentlichen Mechanismen der Teilreaktionen, jedoch ist durch weitere spezielle Experimente in einfachen Fällen auch hierzu eine Aufklärung möglich.

7.3.3.2.1. *Modifikationswechsel reiner Stoffe.* Mehrere Grenzfälle lassen sich für den Mechanismus der Umwandlung diskutieren. Bei den im allgemeinen nur langsam ablaufenden *rekonstruktiven Umwandlungen* entsteht die neue Modifikation über ein Stadium der *Keimbildung*. Mehrere Teilchen lagern sich innerhalb der alten Phase, meist an energetisch bevorzugten Defektstellen, zusammen. Ein so entstehender Keim ist wegen seiner ziemlich großen Oberfläche aber zunächst relativ energiereich und thermodynamisch nicht stabil. Ab einer gewissen Keimgröße wird der Quotient Oberfläche/Volumen aber hinreichend klein, so daß der Keim mit endlicher, meist durch die Andiffusion neuer Teilchen an die Grenzfläche bestimmter Geschwindigkeit zu dem noch energieärmeren Kristall weiterwächst.

Wie bei anderen heterogenen Reaktionen, die mit Phasenneubildung verknüpft sind (s. a. 6.4.5.2.), beobachtet man besonders bei sehr reinen Stoffen Reaktionshemmungen, da erst Keime der neuen Phase entstanden sein müssen. Ein instruktives Beispiel hierfür ist die nur langsame Eintrübung von aus der Schmelze monoklin erstarrtem Schwefel nach Unterschreiten der Umwandlungstemperatur von 368,6 K; sie wird durch die Umbildung zu den rhombischen Kriställchen der Tieftemperaturmodifikation bewirkt. Weitere rekonstruktive Umwandlungen sind bei Calciumcarbonat der Übergang Aragonit → Calcit, bei Zinksulfid der von Zinkblende in Wurtzit.

Rasch, gelegentlich mit Schallgeschwindigkeit, können Modifikationswechsel ablaufen, wenn für die Umwandlung nur kleine Änderungen der Gitterpositionen erforderlich sind. So geht Caesiumchlorid bei 445 °C von dem ihm eigenen CsCl-Typ mit Koordinationszahl 8 durch Git-

terdilatation in [111]-Richtung in eine Modifikation mit NaCl-Struktur über. Ähnlich leicht wie diese *dilative Umwandlung* vollzieht sich die Deformations- oder *displazive Umwandlung* von β- in α-Quarz bei 573 °C; sie ist das Resultat einer Drehung von SiO_4-Tetraedern ohne Zerstörung von Si—O-Bindungen. Die Umwandlungen in andere SiO_2-Modifikationen wie Tridymit oder Cristobalit sind dagegen vom rekonstruktiven Typ, erfordern hohe Aktivierungsenergien und sind demzufolge erschwert.
Zur Unterscheidung der Modifikationswechsel nach ihrer Reversibilität s. a. 6.4.5.2.

7.3.3.2.2. *Zersetzungsreaktionen.* Bedeutender als die schon behandelten *peritektischen Reaktionen* (↑ 6.4.5.4.) mit Zersetzung einer festen in eine neue feste und eine schmelzflüssige Phase sind die, bei denen als Reaktionsprodukte eine feste und eine gasförmige Phase entstehen. Eine mit Hilfe der *Thermogravimetrie* besonders gut zu verfolgende Reaktion ist die thermische Zersetzung von Calciumoxalat-monohydrat,

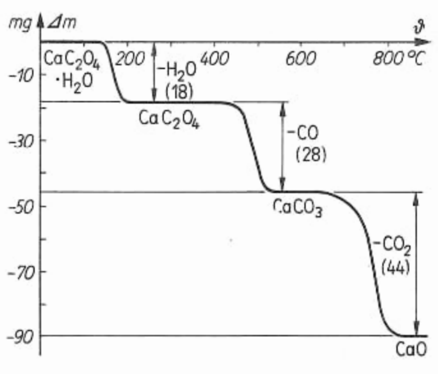


Abb. 7.3.3-1 Masseänderung Δm in Abhängigkeit von der Temperatur ϑ beim thermogravimetrischen Abbau von 1 mmol Calciumoxalat-monohydrat (146 mg) zu Calciumoxid (56 mg) in Stickstoffatmosphäre

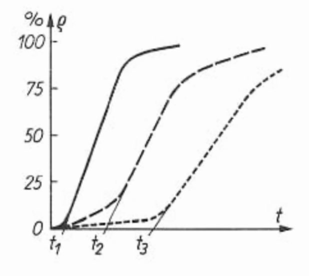


Abb. 7.3.3-2 Abhängigkeit des Reaktionsgrads ϱ von der Zeit t für Festkörperumwandlungen mit unterschiedlichen Induktionsperioden t_i (schematisch)

$CaC_2O_4 \cdot H_2O$ (Abb. 7.3.3-1), weitere Beispiele sind die Entwässerung von Silicaten, die Zersetzung von Alkalimetallnitraten zu Sauerstoff und den entsprechenden Nitriten sowie die fotochemische Spaltung von Silberhalogeniden gemäß $AgX \rightarrow Ag + 1/2\,X_2$.
Wie bei rekonstruktiven Umwandlungen setzt die Zersetzung des alten und die Bildung von Keimen des neuen Feststoffs bevorzugt an Defekten des Kristallgitters ein. Die Form der Umsatz-Zeit-Kurven (Abb. 7.3.3-2) wird stark durch die Dauer der *Induktionsperiode* bestimmt, in der sich die Bildung der Keime vollzieht, die dann im Sinne einer *Autokatalyse* den nachfolgenden Prozeß beschleunigen. Die Induktionsperiode ihrerseits hängt stark von der Realstruktur des Ausgangsstoffs, dessen Vermahlungsgrad und Schüttdichte sowie von der Heizgeschwindigkeit und anderen experimentellen Bedingungen ab. Es bereitet folglich große Schwierigkeiten, stoffspezifische, für die Reaktion des sogenannten normalen Zustands des Festkörpers geltende kinetische Daten, also Geschwindigkeitskonstanten und Aktivierungsenergien, anzugeben.
Die starke Abhängigkeit der Eigenschaften der entstehenden Feststoffe von den Reaktionsbedingungen läßt sich am Gips zeigen. *Gips*, $CaSO_4 \cdot 2\,H_2O$, spaltet schon bei etwa 100 °C einen Teil des Kristallwassers ab unter Bildung des Hemihydrats $CaSO_4 \cdot 1/2\,H_2O$ [*hemisys* griech., halb]. Dieses Produkt wie auch das bis 200 °C daraus entstehende wasserfreie Salz, der *Anhydrit*, bilden wegen ihrer stark gestörten Struktur bei Wasserzugabe rasch Gipskristalle zurück, die miteinander verwachsen und so den Zusammenhalt der erhärtenden Masse bewirken. Bei höheren Temperaturen heilen die Gitterstörungen durch Oberflächendiffusion, Zusammenbacken, Volumendiffusion und Sintern bei gleichzeitiger *Rekristallisation* kleiner Kristalle zu größeren zunehmend aus, und die Abbindefähigkeit mit Wasser geht zurück. Durch Erhitzen auf über 1000 °C entsteht schließlich „totgebrannter Gips"; auch ein solches Calciumsulfat kann aber für die Umsetzung mit Wasser mechanisch aktiviert werden, z. B. durch äußerst feines Vermahlen bis herunter zu Korngrößen im Mikrometerbereich.
Die bei der Gipsentwässerung zu beobachtende Reaktionsbeschleunigung bei oder oberhalb von Zersetzungstemperaturen kann als Beispiel für die *Hedvallregel* angesehen werden. Sie drückt den Sachverhalt aus, daß in Umwandlung begriffene Phasen eine ausgeprägte Reaktionsfreudigkeit haben; die alte Struktur ist weitgehend abgebaut, die neue erst fragmentarisch ausgebildet, so daß zahlreiche reaktive Zentren vorliegen. Besonders deutlich tritt der Hedvalleffekt bei Modifikationswechseln fester Phasen in Erscheinung, wie Abb. 7.3.3-3 für die Reak-

tion von Siliciumdioxid mit Nickel(II)-oxid bzw. Cobalt(II, III)-oxid zeigt. Die Pfeile markieren dort die Temperaturen der Umwandlung von β- (Tief-) in α- (Hoch-) Quarz bzw. von Hochquarz in α- (Hoch-) Tridymit bzw. α- Cristobalit.

7.3.3.2.3. *Reaktionen zwischen festen Phasen*, für die soeben ein Beispiel erwähnt wurde, hielt man bis in das 20. Jahrhundert hinein für unwahrscheinlich, gemäß dem Erfahrungssatz „*Corpora non agunt nisi fluida*" [lat., Körper wirken nicht, wenn sie nicht fließen]. Inzwischen weiß man um die Beweglichkeit von Gitterbausteinen schon weit unterhalb der Schmelztemperaturen der betreffenden Phasen. Ob allerdings aus jeder Phase Teilchen in die jeweils angrenzende übertreten und welche das sind, ist oft nur schwer zu entscheiden. Häufig steht darüber hinaus die Frage, ob nicht auch die über den Feststoffen ausgebildete oder schon vorgegebene Gasphase für den Reaktionsmechanismus wesentlich ist.

Bei der Diffusion geladener Teilchen ist stets die *Elektroneutralitätsbedingung* zu beachten. Die Diffusion von Kationen in die angrenzende Phase muß somit entweder von einem entsprechenden Transport negativ geladener Teilchen, also von Anionen oder von Elektronen, begleitet sein oder durch eine gegenläufige Diffusion von Kationen aus der anderen Phase kompensiert werden. Prinzipiell wären noch weitere Prozesse denkbar, wie die gegenläufige Diffusion negativer Ladungen, die aber – offensichtlich wegen der zu hohen Aktivierungsenergien – gegenüber den schon genannten nicht zum Tragen kommen. Das Eindiffundieren von Teilchen ist im allgemeinen an das Vorliegen entsprechender Leerstellen im Wirtsgitter geknüpft, so daß z. B. der Platzwechsel einer größeren Anzahl von Kationen, der nacheinander durch eine Leerstelle ausgelöst wird (s. a. 6.4.8.2.), auch als gegenläufige Wanderung dieses einen Defekts beschrieben werden kann.

Ein schon klassisches Beispiel für eine additive Reaktion zwischen zwei festen Phasen ist die Bildung des *Spinells*, $MgAl_2O_4$, aus Magnesium- und Aluminiumoxid. Der Mechanismus der Spinellbildung kann dadurch aufgeklärt werden, daß man chemisch weitgehend inerte Markierungen, z. B. feinste Platindrähte, so zwischen zwei in Kontakt kommende Tabletten von MgO und Al_2O_3 befestigt, daß nach Abbruch des Versuchs aus der Lage dieser Markierungen relativ zur Schicht des Reaktionsprodukts auf die Teilchenwanderung geschlossen werden kann (Abb. 7.3.3-4). Bei einseitiger Wanderung eines der Stoffe (AO bzw. M_2O_3 in Abb. 7.3.3-4b, c) in Form seiner Bestandteile ist der Befund nicht eindeutig. Er gestattet keine Entscheidung zwischen einer Wanderung von Kationen und Oxidionen durch die Produktschicht und einer Wanderung von Kationen und Elektronen bei gleichzeitiger Sauerstoffübertragung über die Gasphase. Hierzu sind weitere Informationen einzuholen, z. B. über den Transportverlauf bei unterschiedlichen Partialdrücken an Sauerstoff oder über die *Selbstdiffusionskoeffizienten* und *Überführungszahlen* der einzelnen Bestandteile.

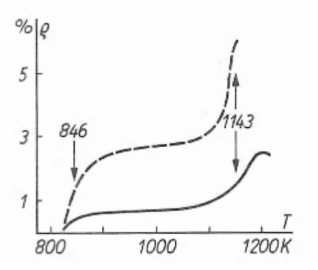


Abb. 7.3.3-3 Reaktionsgrad ϱ der Umsetzung von NiO (*durchgezogene Kurve*) bzw. Co_3O_4 (*gestrichelte Kurve*) mit SiO_2 in Abhängigkeit von der Temperatur T nach HEDVALL

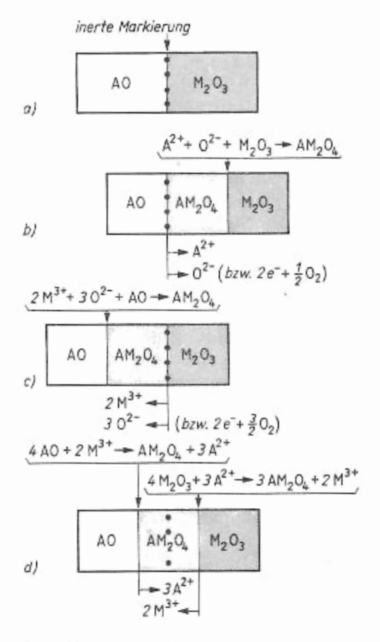


Abb. 7.3.3-4 Schema der Möglichkeiten für den Stofftransport bei der Bildung von Spinellen AM_2O_4, die zugehörigen Lagen der inerten Markierungen Ma und die jeweiligen Phasengrenzreaktionen; a) Tabletten AO/M_2O_3 vor der Umsetzung, b) und c) gleichgerichtete Diffusion von Kationen und negativen Teilchen, d) gegenläufige Kationendiffusion

Die exakte Bestimmung der Lage der Markierungen ist oft nicht möglich, da es zu deren unkontrollierter Verschiebung, z. B. durch die beim Wachstum der Produktschicht auftretenden Spannungen, kommen kann. Eine Variante der Methode arbeitet deshalb mit unterschiedlich breiten Tabletten, die aufeinandergelegt und nach Ablauf der Reaktion bezüglich Lage, Zusammensetzung und Dicke der Produktschicht charakterisiert werden, z. B. mit Hilfe einer *Elektronenstrahlmikrosonde*.

Für die Bildung des eigentlichen Spinells, $MgAl_2O_4$, kann der Grenzfall der gegenläufigen Kationendiffusion (Abb. 7.3.3-4d) als gesichert gelten. In anderen Fällen, z. B. bei der Bildung des Cobalt(II)-chrom(III)-spinells, $CoCr_2O_4$, wird ein solcher Mechanismus nur bei Durchführung der Reaktion in Stickstoffatmosphäre beobachtet. Bei Gegenwart von Sauerstoff diffundieren dagegen Co(II)-ionen über Kationen- und Elektronen über Anionenleerstellen durch die Produktschicht, wobei gleichzeitig Sauerstoff über die Gasphase transportiert wird und für die Reaktion gemäß $Co^{2+} + 2e^- + Cr_2O_3 + 1/2\ O_{2(g)} \rightarrow CoCr_2O_4$ zur Verfügung steht.

Eine gleichgerichtete Diffusion von Kationen und Anionen konnte bei der Spinellbildung bisher nicht gefunden werden, wohl aber beim Zusammentreten einfacher Halogenide zu komplexen, z. B. von KCl und $SrCl_2$ zu K_2SrCl_4. Auf die komplizierten, z. T. ungenügend gesicherten Befunde für die Mechanismen von Silicatbildungsreaktionen, von doppelten Umsetzungen zwischen Halogeniden oder auch von Erdalkalien MO mit Salzen von Oxosäuren, also Carbonaten und Sulfaten, kann hier nicht eingegangen werden.

7.3.3.2.4. *Transportreaktionen.* Bei ihnen wird bei einer Temperatur T_1 ein meist fester Stoff A mit einem Gas B unter Bildung flüchtiger Produkte C umgesetzt, die bei einer Temperatur T_2 durch Zersetzung den Ausgangsstoff A zurückbilden. Die konkreten Bedingungen für einen solchen Stofftransport von einer Temperaturzone in die andere lassen sich bei Kenntnis der thermodynamischen Daten für alle in Frage kommenden Gleichgewichte sowie der Gaspartialdrücke und Diffusionskoeffizienten weitgehend vorhersagen.

Ein bekanntes Beispiel für eine Transportreaktion liegt mit dem *Aufwachsverfahren nach* Van Arkel *und* De Boer (1924) vor, das zur Herstellung zahlreicher Reinstmetalle verwendet werden kann. So bildet sich z. B. aus festem Zirconium und gasförmigem Iod bei 300 bis 600 °C gasförmiges Zirconiumtetraiodid ZrI_4, das an einem dünnen, auf 1500 °C erhitztem Glühdraht zerfällt. Zirconium wird in hochreiner Form abgeschieden, Iod steht für eine erneute Reaktion mit Zirconium im kälteren Teil der Apparatur zur Verfügung. Der meist aus einem anderen Metall, z. B. aus Wolfram, bestehende Glühdraht von einigen Mikrometern Durchmesser spielt im mindestens hundertfach dickeren Metallstab als Verunreinigung kaum eine Rolle.

7.3.3.3. Heterogene Katalyse. Die meisten großtechnischen Verfahren der Synthesechemie schließen heterogen-katalytische Reaktionen ein. Ein *Katalysator* ist nach der Definition von Ostwald (1907) ein Stoff, der schon in sehr geringer Menge die Geschwindigkeit einer gegebenen Reaktion ändert, ohne daß er sich mehr als vorübergehend an der Reaktion beteiligt. Katalysatoren greifen also in die Reaktion ein, indem sie deren Mechanismus verändern, z. B. durch Bildung von Zwischenprodukten die Aktivierungsenergie der Reaktion erniedrigen (s. a. 7.2.4.1.).

So führt die Adsorption von Gasen wie H_2, O_2 oder N_2 an Metalloberflächen zur Ausbildung recht stabiler Bindungen; dadurch werden die ursprünglich in den Gasmolekülen vorhandenen Bindungen gelockert oder gelöst, so daß die Reaktionserleichterung verständlich wird. Wegen der freigesetzten Adsorptionswärme von 80 $kJmol^{-1}$ und mehr und der starken Abhängigkeit der Wechselwirkung zwischen Gas und Unterlage von deren chemischer Natur wird eine solche Adsorption oft als *Chemisorption* bezeichnet und damit von der *Physisorption* abgegrenzt. Bei dieser sind die Wechselwirkungen unspezifisch, so daß man von einer Kondensation des Gases an der als Keim wirkenden Unterlage sprechen kann; bei ihr werden lediglich Energien von 10 bis 40 $kJmol^{-1}$ freigesetzt (s. a. 6.6.4.). Für die heterogene Katalyse werden Festkörper mit möglichst großen Oberflächen und entsprechend zahlreichen Strukturdefekten angestrebt. So entstehen Mischkatalysatoren, indem man auf aktive, hochdispergierte Stoffe, wie Kieselgel, Tonerden und andere Oxide mit spezifischen Oberflächen bis zu 1000 m^2g^{-1}, das katalytisch wirksame Metall in äußerst feiner Verteilung aufbringt. An Luft sind solche Kontakte oft selbstentzündlich, also pyrophor [*pyr* griech., Feuer; *phor* griech., tragen]. Durch die Trägerphase wird eine Kornvergröberung und damit Desaktivierung des Metalls durch Rekristallisation bei der Arbeitstemperatur des Katalysators weitgehend unterbunden. Die hohe Reaktivität bewirkt allerdings, daß sich neben den gewünschten Partnern auch Verunreinigungen des Synthesegases mit dem Katalysator umsetzen können und zu einer *Kontaktvergiftung* führen.

Frühe Beispiele (Döbereiner, um 1820) für heterogene Katalysen waren die leichte Entzündung von Wasserstoff an Platin bei Raumtemperatur in Gegenwart von Luftsauerstoff sowie die Zersetzung von Kaliumchlorat, $KClO_3$, zu Kaliumchlorid und Sauerstoff beim

Erhitzen nach Zusatz von etwas Braunstein, MnO_2. Ohne Braunstein ist die letztere Reaktion zugunsten der Disproportionierung von $KClO_3$ zu Kaliumperchlorat, $KClO_4$, und KCl gehemmt.

Das erste großtechnische Kontaktverfahren (1913) war die *Ammoniaksynthese* nach HABER und BOSCH. Bei Temperaturen um 500 °C und Drücken von 20 bis 50 MPa (200 bis 500 atm) wird Stickstoff sowie Wasserstoff an einem $Fe/K_2O/Al_2O_3$-Mischkatalysator zur Reaktion gebracht. Der Katalysator wird durch Oxydation von reinem Eisen zu Magnetit, Fe_3O_4, hergestellt, wobei die entstehende Schmelze Aluminiumoxid bis zu Massenbrüchen von 3 bis 4% aufnimmt. Bei der nachfolgenden Reduktion des Fe(II,III)-oxids zum Metall bleibt Al_2O_3 unverändert und verhindert durch seine Verteilung zwischen den Eisenkristalliten deren Sinterung. Ein geringer Alkalizusatz von 1 bis 2% erhöht die spezifische Aktivität des Katalysators.

Als geschwindigkeitsbestimmender Schritt der Reaktionsfolge wurde die Adsorption von Stickstoff unter gleichzeitiger Aktivierung der Moleküle an (111)-Flächen von Eisen erkannt. Es ist deshalb wichtig, den Katalysator so vorzubereiten, daß möglichst günstige Bedingungen für die Ausbildung der sonst gegenüber den (100)- und (110)-Flächen zurücktretenden (111)-Flächen des Eisens gegeben sind. Das geschieht durch Herstellen des Katalysators in einem mit dem Synthesegas übereinstimmenden N_2/H_2-Strom. Durch Reaktion mit gleichfalls adsorbiertem Wasserstoff bilden sich dann aller Wahrscheinlichkeit nach adsorbierte N_2H-Gruppen, die stufenweise durch Umsetzung mit adsorbiertem oder molekularem Wasserstoff in Ammoniak übergehen.

7.3.4. Elektrochemische Reaktionen

Am Aufbau der Materie sind in allen Aggregatzuständen elektrische Ladungsträger beteiligt. Sie sind häufig frei beweglich, wie die Elektronen in Metallen oder Halbleitern oder die Ionen in Salzlösungen oder -schmelzen. Diese Ladungsträger können die Strukturen an Phasengrenzen in heterogenen stofflichen Systemen stark beeinflussen und durch die Phasengrenzen hindurchtreten. Dabei fließt ein Strom, der vielfach mit elektrochemischen Reaktionen verbunden ist, die zu chemischen Stoffumsätzen führen. Die Untersuchung und Beschreibung von *Gleichgewichtszuständen* elektrochemischer Reaktionen gehört in das Gebiet der *elektrochemischen Thermodynamik*, *zeitliche Veränderungen* in elektrochemischen Systemen sind Gegenstand der *elektrochemischen Kinetik*.

7.3.4.1. Grundbegriffe

7.3.4.1.1. *Elektrochemische Reaktionen und Elektroden.* Schließt man z. B. zwei Zinkbleche, die in eine wäßrige Zinksulfatlösung tauchen, an die Pole einer Gleichspannungsquelle an, so fließt ein Strom, der mit stofflichen Umwandlungen an den Blechen verbunden ist. Am positiv gepolten Blech treten positiv geladene Zinkionen Zn^{2+} aus dem Metall in die Lösung, am negativ gepolten Blech umgekehrt aus der Lösung in das Metall über. Der Zinkionenübergang stellt einen einfachen Fall einer elektrochemischen Reaktion dar.

Elektrochemische Reaktionen sind allgemein heterogene Reaktionen, die an Grenzflächen benachbarter, elektrisch leitender Phasen stattfinden und mit dem Durchtritt von elektrischen Ladungsträgern durch die Grenzfläche verbunden sind. Mindestens ein Reaktionsteilnehmer muß dabei ein Ion sein.

Das Zweiphasensystem $Zn(I)/Zn^{2+} \cdot aq(II)$ aus dem metallischen Zink und der wäßrigen Zinksalzlösung, im täglichen Sprachgebrauch allerdings auch das Zinkblech allein, wird als *Elektrode* bezeichnet. Bei *Metallionenelektroden* ist der *durchtrittsfähige Ladungsträger* ein Metallkation, z. B. Zn^{2+}. Die elektrochemische Reaktion läßt sich dann durch die Bruttoreaktionsgleichung $Zn^{2+}(I) \rightleftharpoons Zn^{2+}(II)$ oder $Zn(I) \rightleftharpoons Zn^{2+}(II) + 2e^-(I)$ beschreiben. I ist hier das Metall, II die Lösung (Abb. 7.3.4-1). Bei der Silber-Silberchlorid-Elektrode $Ag(I)/AgCl(II)/Cl^- \cdot aq(III)$ ist das Silber mit einer Deckschicht von festem Silberchlorid überzogen. Bei solchen *Elektroden zweiter Art* erfolgt der Ladungsaustausch zwischen dem Metall und der Lösung über eine feste Verbindung, hier das Silberchlorid. Dabei ist ein Anion, das negativ geladene Chloridion, der durchtrittsfähige Ladungsträger in der elektrochemischen Reaktion $Cl^-(II) \rightleftharpoons Cl^-(III)$. Elektroden, bei denen ein Anion durch die Phasengrenze tritt, bezeichnet man als *Anionenelektroden*. Zum Unterschied

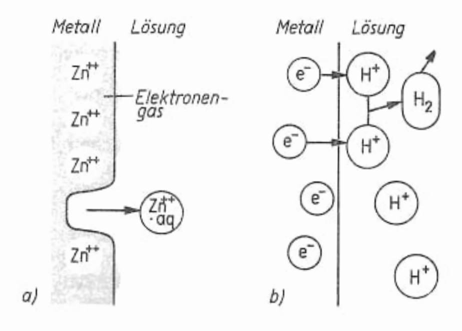


Abb. 7.3.4-1 a) Zinkelektrode (Metallionenelektrode), b) Wasserstoffelektrode (Redoxelektrode)

davon ist bei *Redoxelektroden* aus einem inerten Metall und der Lösung eines Redoxsystems das Elektron der durchtrittsfähige Ladungsträger. Beispiele dafür sind die Wasserstoffelektrode $Pt(I)/H_2$, $H^+ \cdot aq(II)$ (Abb. 7.3.4-1 b), die Chlorelektrode $Pt(I)/Cl_2$, $Cl^- \cdot aq(II)$ und die Eisen(2)-Eisen(3)-Elektrode $Pt(I)/Fe^{2+} \cdot aq$, $Fe^{3+} \cdot aq(II)$ mit den elektrochemischen Reaktionen $H_2(II) \rightleftharpoons 2\,H^+(II) + 2e^-(I)$, $2\,Cl^-(II) \rightleftharpoons Cl_2(II) + 2e^-(I)$ und $Fe^{2+}(II) \rightleftharpoons Fe^{3+}(II) + e^-(I)$.

7.3.4.1.2. *Faradaysche Gesetze.* Die Faradayschen Gesetze geben den quantitativen Zusammenhang zwischen Ladungsmenge und Stoffumsatz bei elektrochemischen Reaktionen an. Das *erste Faradaysche Gesetz* besagt, daß der Stoffumsatz der durch eine Elektrode geflossenen Elektrizitätsmenge proportional ist. Nach dem *zweiten Faradayschen Gesetz* verhalten sich die durch gleiche Elektrizitätsmengen erzielten Umsätze verschiedener Stoffe wie deren Äquivalentmassen.

Beide Gesetze lassen sich in der Beziehung $m_i = \ddot{A}_i q/F$ zusammenfassen. Dabei ist m_i die umgesetzte Masse des Stoffs i, $q = \int I\,dt$ die Elektrizitätsmenge, $F = 96487$ As mol^{-1} die *Faraday-konstante* (s. a. 6.3.4.) und $\ddot{A}_i$ die Äquivalentmasse des umgesetzten Stoffs i in der betrachteten elektrochemischen Reaktion.

Eine direkte Anwendung erfahren die Faradayschen Gesetze bei der *coulometrischen Analyse* zur Bestimmung kleiner Substanzmengen. Läuft z. B. an einer Elektrode nur die Reaktion $Cu^{2+} \cdot aq + 2e^- \rightarrow Cu$ ab, bis keine Kupferionen mehr in der Lösung vorhanden sind, so läßt sich aus der geflossenen Elektrizitätsmenge die abgeschiedene Kupfermenge berechnen. Die Äquivalentmasse des Kupfers ist bei dieser Reaktion $\ddot{A}_{Cu} = A_{Cu}/2 = 31{,}77$ gmol^{-1}, da zwei Faraday zur Entladung von einem Mol Kupferionen notwendig sind. A_{Cu} ist die Atommasse des Kupfers. Da auch kleine Elektrizitätsmengen einfach und genau bestimmbar sind, wird die coulometrische Analyse bevorzugt für geringe Substanzmengen angewandt. Umgekehrt lassen sich auch Elektrizitätsmengen mit *Coulometern*, wie Silber-, Knallgas- oder Iodcoulometern, messen.

7.3.4.1.3. *Galvanispannung einer Elektrode.* Sie ist die elektrische Potentialdifferenz zwischen dem Inneren der beiden benachbarten Phasen einer Elektrode. Taucht man ein Zinkblech in eine wäßrige Lösung von Zinksulfat, so treten positiv geladene Zinkionen aus dem Metall in die Lösung über. Dabei lädt sich das Metall negativ gegenüber der Lösung auf, weil jedes Zinkion zwei Elektronen in dem Metall hinterläßt, während die Lösung zusätzlich positive Überschußladungen erhält. Dieser Vorgang kommt aber nach sehr kleinem, durch chemische Analyse nicht nachweisbarem Umsatz zum Stillstand und führt zu einem Gleichgewicht, weil die entstehende elektrische Potentialdifferenz den Durchtritt weiterer Ladungsträger verhindert. Ein anderer Anteil der Galvanispannung kommt durch die Ausbildung von Dipolschichten an der Phasengrenze zustande. An einer Metalloberfläche treten die Elektronen infolge ihrer kinetischen Energie etwas weiter als die um ihre Ruhelage schwingenden Metallionen hervor und bilden so mit diesen eine Dipolschicht aus. Auch Moleküle mit Dipolstruktur, z. B. Wassermoleküle, haben in Grenzflächen häufig Vorzugsrichtungen, die elektrische Potentialdifferenzen bedingen. Für die Galvanispannung zwischen zwei Phasen I und II gilt $g^{I,II} = \varphi^I - \varphi^{II} = \psi^{I,II} + \chi^{I,II}$. $\psi^{I,II}$ ist die *Voltaspannung*, verursacht durch Überschußladungen, und $\chi^{I,II}$ die *Oberflächenpotentialdifferenz* infolge der Dipolschichten.

Die Größe der *Gleichgewichtsgalvanispannung* $g_{gl}^{I,II}$ einer *einfachen Elektrode*, an der nur *eine* elektrochemische Reaktion abläuft, hängt von der Lage des Gleichgewichts der Reaktion ab. Für eine Reaktion $\nu_A A + \nu_B B \rightleftharpoons \nu_C C + \nu_D D$ zwischen neutralen Teilchen gilt im Gleichgewicht $\nu_A \mu_A + \nu_B \mu_B = \nu_C \mu_C + \nu_D \mu_D$ (s. a. 6.1.2.4.3.). Bei elektrisch geladenen Teilchen wie Ionen oder Elektronen ist noch die elektrische Arbeit $z_i F g^{I,II}$ beim Übergang von einem Mol Ladungsträger mit der Ladungszahl z_i aus der Phase I in die Phase II zu berücksichtigen. Deshalb muß zu dem chemischen Potential $\mu_i = \mu_i^{\ominus} + RT \ln a_i$ der Teilchen noch ein Term hinzugefügt werden, der ihre elektrische Energie in der jeweiligen Phase α berücksichtigt. Man kommt dann zu dem *elektrochemischen Potential* $\tilde{\mu}_i = \mu_i + z_i F \varphi^{\alpha}$. Für die elektrochemische Reaktion $Zn^{2+}(I) \rightleftharpoons Zn^{2+}(II)$ der Zinkionenelektrode lautet die Gleichgewichtsbedingung $\tilde{\mu}^{I}_{Zn^{2+}} = \tilde{\mu}^{II}_{Zn^{2+}}$, aus der die nach ihrem Entdecker benannte *Nernstsche Gleichung* folgt

$$\varphi^I - \varphi^{II} = g_{gl}^{I,II} = g_{Zn}^{\ominus I,II} + \frac{RT}{2F} \ln \frac{a^{II}_{Zn^{2+}}}{a^{I}_{Zn^{2+}}}$$

Darin bedeutet $\varphi^I - \varphi^{II} = g_{gl}^{I,II}$ die *Gleichgewichtsgalvanispannung*, $g_{Zn}^{\ominus}$ die *Standardgalvanispannung* der Zinkionenelektrode. Für eine beliebige elektrochemische Reaktion $\sum |\nu_j|\, R_j(\alpha) \rightleftharpoons \sum \nu_k R_k(\alpha)$ der Ausgangsstoffe R_j zu den Endstoffen R_k in den Phasen α hat die Nernstsche Gleichung die allgemeine Form $g_{gl}^{I,II} = g^{\ominus} + \frac{RT}{z_r F} \sum_{\alpha, i} \nu_i^{\alpha} \ln a_i^{\alpha}$, worin z_r die *Reaktionsladungszahl* bedeutet. Die stöchiometrischen Koeffizienten ν_j der Ausgangsprodukte sind negativ, und ν_k der Endprodukte positiv. z_r gibt an, wieviel positive Äquivalentladungen F bei einem Formelumsatz von der Phase I nach der Phase II transportiert werden.

Galvanispannungen entstehen an den Grenzen verschiedener Phasen unabhängig von ihren

Aggregatzuständen. Die an *Kolloidteilchen* auftretenden Galvanispannungen sind oftmals für deren Stabilität verantwortlich. Potentialdifferenzen zwischen Elektrolytlösungen unterschiedlicher Zusammensetzung bezeichnet man als *Diffusionsspannungen*. Andere Galvanispannungen mit besonderen Namen sind *Donnanspannungen* zwischen semipermeablen Membranen und Flüssigkeiten, die in der Biologie große Bedeutung haben, und *Kontaktspannungen* zwischen verschiedenen Metallen. Galvanispannungen an Halbleitern sind die Grundlage der Transistortechnik und Mikroelektronik.

7.3.4.1.4. *Galvanische Zellen.* Sie entstehen durch Kombination zweier Elektroden, z. B. Metall(I)/Lösung(II)/Lösung(III)/Metall(IV)/Metall(I'). Die beiden Endphasen (I) und (I') sind die *Pole* der galvanischen Zelle. Die Differenz der inneren elektrischen Potentiale φ^{I} und $\varphi^{\mathrm{I'}}$ wird als Zellspannung $U = \varphi^{\mathrm{I}} - \varphi^{\mathrm{I'}}$ bezeichnet. Fließt kein Strom durch die Zelle, so setzt sich die Zellspannung additiv aus den einzelnen Galvanispannungen zusammen: $U = \varphi^{\mathrm{I}} - \varphi^{\mathrm{I'}} = g^{\mathrm{I,II}} + g^{\mathrm{II,III}} + g^{\mathrm{III,IV}} + g^{\mathrm{IV,I'}}$ (Abb. 7.3.4-2). Sie ist direkt meßbar, wenn die Pole I und I' der Zelle aus dem gleichen Material bestehen. Bei Stromfluß entstehen im Inneren der Phasen zusätzliche Potentialdifferenzen, die häufig durch das Ohmsche Gesetz $U = IR^{\alpha}$ beschreibbar sind, mit der Stromstärke I und dem Ohmschen Widerstand R^{α} der Phase α. Die Galvanispannungen hängen dann auch noch von der Größe des Stromflusses ab (↑ Abb. 7.3.4-2). Die Kombination von zwei einfachen Elektroden

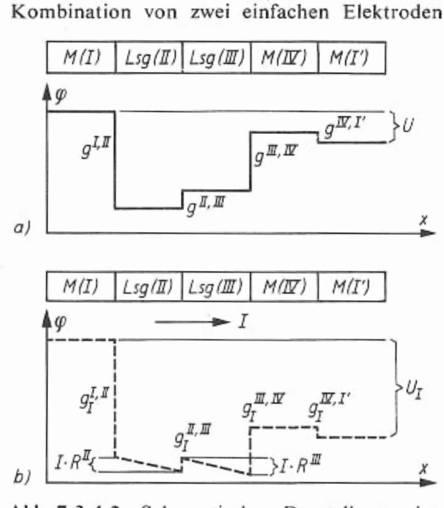


Abb. 7.3.4-2 Schematische Darstellung des Potentialverlaufs φ in einer galvanischen Zelle; a) ohne Stromfluß, b) mit Stromfluß; *M* Metall, *Lsg.* Lösung, *x* Ortskoordinate, *U* Zellspannung, U_I Zellspannung bei Stromfluß, *R* Widerstand, *I* Stromstärke

mit einer gemeinsamen Elektrolytlösung ergibt eine *ideale galvanische Zelle*, solange kein Strom durchfließt. Die *Chlorknallgaszelle* besteht z. B. aus zwei Platinblechen, die in eine Salzsäurelösung tauchen: Pt (I)/H_2, HCl · aq, Cl_2(II)/Pt(I'). Ein Pt-Blech wird von Chlor-, das andere von Wasserstoffgas umspült. Verbindet man die beiden Pole der Zelle über einen Widerstand, so fließt ein Strom. Dabei findet an der Wasserstoffelektrode die elektrochemische Reaktion $H_2(\mathrm{II}) \rightarrow 2\,H^+(\mathrm{II}) + 2e^-(\mathrm{I})$ statt. Die vom Platin aufgenommenen Elektronen fließen über den Widerstand zur Chlorelektrode und werden dort durch die elektrochemische Reaktion $Cl_2(\mathrm{II}) + 2e^-(\mathrm{I'}) \rightarrow 2\,Cl^-(\mathrm{II})$ verbraucht. Insgesamt läuft in der Zelle die rein chemische *Zellreaktion* $H_2 + Cl_2 \rightarrow 2\,H^+ + 2\,Cl^-$ ab, d. h. die Bildung von Chlorwasserstoff aus Wasserstoff und Chlor. Dabei liefert die Zelle elektrische Energie durch chemische Vorgänge, sie wirkt als *galvanisches Element*. Verbindet man andererseits die Wasserstoffelektrode mit dem negativen, die Chlorelektrode mit dem positiven Pol einer *äußeren* Gleichspannungsquelle, so kann man die Richtung der elektrochemischen Reaktion an den beiden Elektroden und damit die Richtung der Zellreaktion umkehren: $2\,H^+ + 2\,Cl^- \rightarrow H_2 + Cl_2$. Elektroenergie wird verbraucht und in chemischen Produkten mit höherem Energieinhalt im Vergleich zu den Ausgangsstoffen gespeichert. Dieser durch äußere Spannung erzwungene Vorgang wird als *Elektrolyse* bezeichnet. Fließt kein Strom durch die Zelle und befinden sich die beiden Elektroden im Gleichgewicht, so ist die reversible Zellspannung $U_{\mathrm{gl}} = g_{\mathrm{gl}}^{\mathrm{I,II}} - g_{\mathrm{gl}}^{\mathrm{I',II}}$ gleich der Differenz der Gleichgewichtsgalvanispannungen der Wasserstoff- und Chlorelektrode.

Das *Daniellelement* Zn(I)/Zn^{2+}(II)/Cu^{2+}(III)/Cu(IV)/Zn(I') als Kombination einer Zinkionen- mit einer Kupferionenelektrode wurde früher zur Stromlieferung genutzt. An den Elektroden können die Reaktionen $\mathrm{Zn(I)} \rightleftharpoons Zn^{2+}(\mathrm{II}) + 2e^-(\mathrm{I})$ bzw. $\mathrm{Cu(IV)} \rightleftharpoons Cu^{2+}(\mathrm{III}) + 2e^-(\mathrm{IV})$ ablaufen. Sie ergeben zusammen die Zellreaktion $\mathrm{Zn} + Cu^{2+} \rightleftharpoons Zn^{2+} + \mathrm{Cu}$. An der Berührungsfläche beider Elektrolytlösungen finden Diffusionsvorgänge statt. Durch unterschiedliche Diffusionsgeschwindigkeiten der beteiligten Ionen bildet sich eine elektrische Potentialdifferenz $g^{\mathrm{II,III}} = g_{\mathrm{Diff}}$ zwischen den Lösungen aus, eine Diffusionsspannung. Ihre Größe läßt sich nur für einige einfache Fälle exakt berechnen. Wenn zwei verdünnte Lösungen des gleichen 1,1-wertigen Elektrolyten mit unterschiedlicher Konzentration aneinandergrenzen, ergibt sich z. B. für die Diffusionsspannung $g_{\mathrm{Diff}} = RT/\mathrm{F} \cdot [(\Lambda_{\infty+} - \Lambda_{\infty-})/(\Lambda_{\infty+} + \Lambda_{\infty-})] \cdot [\ln a_{\pm}^{\mathrm{II}}/a_{\pm}^{\mathrm{I}}]$. $a_{\pm}^{\mathrm{I}}$ und $a_{\pm}^{\mathrm{II}}$ sind die mittleren Aktivitäten in den beiden Lösungen I und II, Λ_{∞} ist

die Ionenäquivalentleitfähigkeit bei unendlicher Verdünnung (Index + für Kation, – für Anion). *Zellen mit Diffusionsspannung* zählen stets zu den *realen galvanischen Zellen.* Bei Zellspannungsmessungen kann man Diffusionsspannungen weitgehend unterdrücken, wenn man zwischen die beiden Lösungen der Elektroden eine konzentrierte Salzlösung mit einwertigen Ionen bringt, die möglichst gleiche Äquivalentleitfähigkeiten und damit gleiche Diffusionskoeffizienten haben. Die zwei dabei auftretenden Diffusionsspannungen sind sehr klein und entgegengerichtet, so daß sie sich weitgehend kompensieren. Als Salze werden dafür bevorzugt Kaliumchlorid, Kalium- oder Ammoniumnitrat verwendet.

Differenzen des inneren elektrischen Potentials lassen sich grundsätzlich nur zwischen Phasen gleicher Struktur und Zusammensetzung messen. In allen anderen Fällen treten z. B. bei dem Anschluß an Meßgeräte weitere Galvanispannungen in Form von Kontaktspannungen auf, die sich nicht kompensieren und das Meßergebnis verfälschen. Aus diesem Grund ist die Galvanispannung einer einzelnen Elektrode nicht meßbar. Man kann die betrachtete Elektrode aber mit einer *Bezugselektrode* zu einer galvanischen Zelle kombinieren. Die Spannung dieser *Bezugszelle* bezeichnet man als *Bezugszellspannung*, als *relative Elektrodenspannung* oder kurz als *Elektrodenpotential* der jeweiligen Elektrode. Eventuell auftretende Diffusionsspannungen müssen rechnerisch oder experimentell eliminiert werden. Als Bezugselektrode für den Vergleich der Galvanispannungen verschiedener Elektroden dient meist die *Standardwasserstoffelektrode*, d. h. eine Wasserstoffelektrode unter Standardbedingungen, bei einem Partialdruck des Wasserstoffs $p_{H_2} = 101 \cdot 10^3$ Pa (1 atm) und einer Wasserstoffionenaktivität $a_{H^+} = 1$ in der Lösung (Abb. 7.3.4-3). Die Elektrodenspannung einer Zinkionenelektrode ist dann beispielsweise gleich der Zellspannung. Es gilt $U_H = \varphi^{I} - \varphi^{I'}$ der Zelle $Zn(I)/Zn^{2+} \cdot aq(II)/H^+ \cdot aq\,(a_{H^+} = 1), H_2\,(p_{H_2} = 101 \cdot 10^3\ Pa)\,(III)/Pt(IV)/\,Zn(I')$ abzüglich der Diffusionsspannung $g^{II,III}$.

Befindet sich die Zinkionenelektrode ebenfalls im Standardzustand, d. h., besteht sie aus reinem Zink und einer Zinkionenlösung mit der

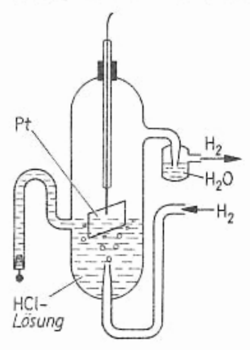


Abb. 7.3.4-3 Schematischer Aufbau einer Wasserstoffelektrode

Tab. 7.3.4-4 Standardbezugszellspannungen $U_H^{\ominus}$ in V von Metallionenelektroden *ME* in wäßrigen Lösungen bei 25 °C

ME	$U_H^{\ominus}$	*ME*	$U_H^{\ominus}$	*ME*	$U_H^{\ominus}$
Li/Li^+	−3,045	Dy/Dy^{3+}	−2,353	Cr/Cr^{3+}	−0,744
K/K^+	−2,925	Am/Am^{3+}	−2,320	Ga/Ga^{3+}	−0,529
Rb/Rb^+	−2,925	Ho/Ho^{3+}	−2,319	Fe/Fe^{2+}	−0,440
Cs/Cs^+	−2,923	Er/Er^{3+}	−2,296	Cd/Cd^{2+}	−0,403
Ra/Ra^+	−2,916	Tm/Tm^{3+}	−2,278	In/In^{3+}	−0,343
Ba/Ba^{2+}	−2,906	Yb/Yb^{3+}	−2,267	Tl/Tl^+	−0,336
Sr/Sr^{2+}	−2,888	Lu/Lu^{3+}	−2,255	Co/Co^{2+}	−0,277
Ca/Ca^{2+}	−2,866	Sc/Sc^{3+}	−2,077	Ni/Ni^{2+}	−0,250
Na/Na^+	−2,714	Pu/Pu^{3+}	−2,031	Sn/Sn^{2+}	−0,136
La/La^{3+}	−2,522	Th/Th^{4+}	−1,899	Pb/Pb^{2+}	−0,126
Ce/Ce^{3+}	−2,483	Np/Np^{3+}	−1,856	H_2/H^+	0
Pr/Pr^{3+}	−2,462	Be/Be^{2+}	−1,847	Cu/Cu^{2+}	+0,342
Nd/Nd^{3+}	−2,431	U/U^{3+}	−1,789	Cu/Cu^+	+0,521
Pm/Pm^{3+}	−2,423	Al/Al^{3+}	−1,662	Hg/Hg^{2+}	+0,788
Sm/Sm^{3+}	−2,414	Ti/Ti^{2+}	−1,628	Ag/Ag^+	+0,799
Eu/Eu^{3+}	−2,407	Zr/Zr^{4+}	−1,529	Rh/Rh^{3+}	+0,80
Gd/Gd^{3+}	−2,397	V/V^{2+}	−1,186	Pd/Pd^{2+}	+0,987
Tb/Tb^{3+}	−2,391	Mn/Mn^{2+}	−1,180	Pt/Pt^{3+}	+1,2
Y/Y^{3+}	−2,372	Nb/Nb^{3+}	−1,099	Au/Au^{3+}	+1,498
Mg/Mg^{2+}	−2,363	Zn/Zn^{2+}	−0,763	Au/Au^+	+1,691

Aktivität $a_{Zn^{2+}} = 1$, so spricht man von der *Standardbezugszellspannung* oder vom *Standardelektrodenpotential* $U_H^{\ominus}(Zn/Zn^{2+})$. Ordnet man alle möglichen Metallionenelektroden nach ihren Standardbezugszellspannungen $U_H^{\ominus}$ in einer Reihe an, so ergibt sich ein Bereich von etwa -3 V bei der Lithiumelektrode bis zu etwa $+1{,}5$ V bei der Goldelektrode. Diese Reihe wird als *Spannungsreihe der Metalle* bezeichnet. Die *unedlen* Metalle haben negative, die *edlen* positive Standardbezugszellspannungen. Dadurch können aus der Stellung in der Spannungsreihe bestimmte Schlüsse auf die chemische Beständigkeit der Metalle gezogen werden. Die Bezugszellspannungen für Metallionenelektroden sind in Tabelle 7.3.4-4 für 25 °C zusammengestellt. Eine ähnliche Spannungsreihe läßt sich für Redoxelektroden aufstellen (Tab. 7.3.4-5). Starke *Oxydationsmittel* haben positive, starke *Reduktionsmittel* negative Standardbezugszellspannungen.

Anstelle der Wasserstoffelektrode verwendet man in der Praxis verschiedene bequemer zu handhabende Bezugselektroden. Besonders gut reproduzierbar sind *Elektroden zweiter Art*, bei denen mit der Lösung ein fester Bodenkörper im Gleichgewicht steht oder das Elektrodenmetall mit einer Deckschicht überzogen ist. Bezugszellspannungen verschiedener solcher Elektroden sind in Tab. 7.3.4-6 ersichtlich. Typische Vertreter dieser Art sind die Silber-Silberchlorid-Elektrode $Ag/AgCl(fest)/Cl^- \cdot aq$ oder die Quecksilber-Quecksilber(I)-chlorid-Elektrode, auch Kalomelelektrode genannt, $Hg/Hg_2Cl_2(fest)/Cl^- \cdot aq$ (Abb. 7.3.4-7). Bei diesen Elektroden bestimmt die Aktivität des Anions des Bodenkörpers die Galvanispannung. So lautet die Nernstsche Gleichung für die Silber-Silberchlorid-Elektrode $U = U_{AgCl}^{\ominus} - (RT/F)\ln a_{Cl^-}$. In einer idealen galvanischen Zelle, die z. B. den Aufbau Ag(I)/AgCl (gesättigt), $HCl \cdot aq$, H_2(II)/Pt(III)/Ag(I′) hat, läuft die Zellreaktion $2\,Ag(I) + 2\,H^+(II) + 2\,Cl^-(II) \rightleftharpoons 2\,AgCl(II) + H_2(II)$ ab. Die damit verbundene freie Reaktionsenthalpie $\Delta_R G$ (6.1.2.) hängt mit der Zellspannung U_{gl} einer idealen galvanischen Zelle im Gleichgewicht über die Beziehung $\Delta_R G = z_r U_{gl} F$ zusammen. Die Messung der Gleichgewichtszellspannung gestattet somit die Ermittlung von $\Delta_R G$-Werten und anderen thermodynamischen Daten mit großer Präzision, wenn sich für die jeweiligen Zellreaktionen geeignete ideale galvanische Zellen aufbauen

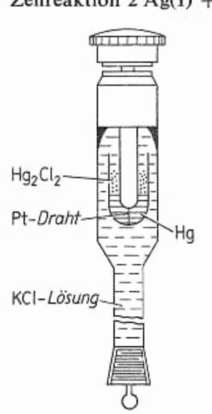


Abb. 7.3.4-7 Schematischer Aufbau einer $Hg/Hg_2Cl_2/KCl^- \cdot aq$-Bezugselektrode (Kalomelelektrode)

Tab. 7.3.4-5 Standardbezugszellspannungen $U_H^{\ominus}$ in *V* einiger Redoxsysteme *RS* in wäßriger Lösung für 25 °C

RS	$U_H^{\ominus}$
Cr^{2+}/Cr^{3+}	$-0{,}408$
Ti^{2+}/Ti^{3+}	$-0{,}369$
V^{2+}/V^{3+}	$-0{,}256$
H_2/H^+	0
Sn^{2+}/Sn^{4+}	$+0{,}15$
Cu^+/Cu^{2+}	$+0{,}153$
$Fe(CN)_6^{4-}/Fe(CN)_6^{3-}$	$+0{,}36$
I_2/I^-	$+0{,}536$
H_2O_2/O_2	$+0{,}682$
Chinon/Hydrochinon	$+0{,}699$
Fe^{2+}/Fe^{3+}	$+0{,}771$
Hg_2^{2+}/Hg^{2+}	$+0{,}920$
Br_2/Br^-	$+1{,}065$
O_2/H_2O	$+1{,}229$
Tl^+/Tl^{3+}	$+1{,}25$
Cl_2/Cl^-	$+1{,}360$
Mn^{2+}/Mn^{3+}	$+1{,}51$
Ce^{3+}/Ce^{4+}	$+1{,}61$
Co^{2+}/Co^{3+}	$+1{,}82$
Ag^+/Ag^{2+}	$+1{,}98$
F_2/F^-	$+2{,}87$

Tab. 7.3.4-6 Bezugszellspannung U_H in V von häufig verwendeten Bezugselektroden *BE* für 25 °C

BE	U_H
$Hg/Hg_2Cl_2/KCl$ ($a = 1$)	$+0{,}2679$
$Hg/Hg_2Cl_2/KCl$ (gesättigt)	$+0{,}241$
$Hg/Hg_2Cl_2/KCl$ (1n)	$+0{,}280$
$Hg/Hg_2Cl_2/KCl$ (0,1n)	$+0{,}334$
$Hg/Hg_2Cl_2/KCl$ (0,01n)	$+0{,}388$
$Hg/Hg_2SO_4/H_2SO_4$ ($a = 1$)	$+0{,}6141$
$Hg/Hg_2SO_4/H_2SO_4$ (0,1n)	$+0{,}705$
$Ag/AgCl/KCl$ ($a = 1$)	$+0{,}2224$
$Ag/AgCl/KCl$ (gesättigt)	$+0{,}197$
$Ag/AgCl/KCl$ (1n)	$+0{,}236$
$Ag/AgCl/KCl$ (0,1n)	$+0{,}290$
$Ag/AgCl/KCl$ (0,01n)	$+0{,}343$
$Hg/HgO/KOH$ ($a = 1$)	$+0{,}0984$
$Hg/HgO/KOH$ (1n)	$+0{,}107$
$Hg/HgO/KOH$ (0,1n)	$+0{,}162$

lassen. Als weitere thermodynamische Größe kann man die Entropieänderung $\Delta_R S$ einer Zellreaktion aus der *Temperaturabhängigkeit der Gleichgewichtszellspannung* nach $\Delta_R S = -z_r F\left(\frac{\partial U_{gl}}{\partial T}\right)_p$ ermitteln. Der *Druckkoeffizient der Gleichgewichtszellspannung* steht mit der Volumenänderung $\Delta_R V$ bei der Zellreaktion im Zusammenhang $\Delta_R V = z_r F\left(\frac{\partial U_{gl}}{\partial p}\right)_T$. Er nimmt allerdings nur dann meßbare Werte an, wenn im Verlaufe der Reaktion Gase gebildet oder verbraucht werden. Bei reversibel und isotherm-isobar ablaufender Zellreaktion gilt für die Wärmetönung aufgrund der Gibbs-Helmholtzschen Beziehungen $\Delta Q_{rev} = T\Delta_R S = \Delta_R H - \Delta_R G$. Sie wird als *Peltierwärme* bezeichnet und unterscheidet sich von der Wärmetönung $(\Delta Q)_p = \Delta_R H$ einer chemischen Reaktion, die durch die gleiche Bruttoreaktionsgleichung beschrieben wird, aber nicht in einer galvanischen Zelle abläuft, um den Betrag $\Delta_R G$. Außer der Zellspannung wird noch die *chemische Spannung E* verwendet. Sie soll die ältere Bezeichnung *elektromotorische Kraft* (EMK) ersetzen. Zur Gleichgewichtszellspannung besteht der Zusammenhang $U_{gl} = -E$.

7.3.4.1.5. *Elektrochemische Doppelschicht.* Die elektrische Potentialdifferenz zwischen den Phasen einer Elektrode hat zwei Ursachen. Eine davon ist der Übertritt von Ladungsträgern aus einer Phase in die andere. Die auf beiden Seiten der Phasengrenze auftretenden Überschußladungen entgegengesetzten Vorzeichens wandern nicht in das Innere der leitenden Phasen ab, sondern bilden aufgrund der elektrostatischen Anziehungskräfte eine *elektrochemische Doppelschicht.* Nach den Vorstellungen von HELMHOLTZ ordnen sich die überschüssigen Ionen bzw. Elektronen in zwei zur Phasengrenze parallelen Schichten an, deren Abstand ungefähr einem Ionenradius entspricht (Abb. 7.3.4-8). Die *Überschußionen auf der Lösungsseite* sind gewöhnlich solvatisiert, d. h. von Lösungsmittelmolekülen umgeben (↑ 6.4.3.). Die Fläche, die man durch ihre Mittelpunkte legen kann, wird *äußere Helmholtzfläche*, der Raum zwischen ihr und der entsprechenden Fläche durch die Ladungsmittelpunkte auf der Metallseite *starre Schicht* oder *äußere Helmholtzschicht* genannt. Dieses Modell ist stark idealisiert, stellt jedoch für gut leitende Phasen, also Metalle und Lösungen mit hoher Elektrolytkonzentration, eine brauchbare Näherung dar. Infolge der Wärmebewegung der Ionen erfolgt vielfach eine Auflockerung dieser Struktur, die zu einer diffusen Verteilung der Überschußionen führt und besonders bei niedrigen Elektrolytkonzentrationen in Erscheinung tritt. Deshalb weist die elektrochemische Doppelschicht einen starren und einen diffusen Bereich auf. Die Anzahl der Überschußionen in der Lösung nimmt mit steigendem Abstand von der Phasengrenze exponentiell ab. Bei verdünnten Lösungen hat die diffuse Schicht eine Dicke in der Größenordnung von 10^{-4} cm. Bei hohen Elektrolytkonzentrationen kommt sie mit 10^{-7} bis 10^{-8} cm in die Größenordnung der starren Schicht.

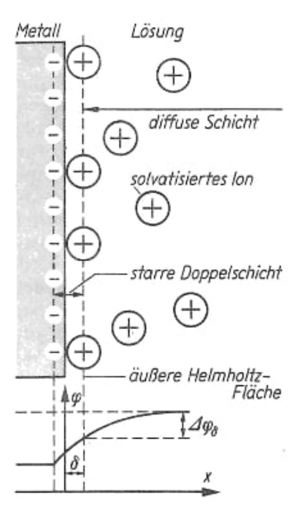


Abb. 7.3.4-8 Schematische Darstellung der Struktur und des Potentialverlaufs in der elektrochemischen Doppelschicht (nach STERN)

In vielen Fällen findet eine *spezifische Adsorption* von Ionen der Lösung, insbesondere von Anionen, an der Metalloberfläche statt. Die spezifisch adsorbierten Ionen werden durch chemische Bindungskräfte so stark angezogen, daß sie auf der Metallseite ihre Solvathüllen verlieren. Dadurch liegen ihre Mittelpunkte näher an dem Metall als die der vollständig solvatisierten Überschußionen. Die Fläche durch die Mittelpunkte der spezifisch adsorbierten Ionen wird als *innere Helmholtzfläche* bezeichnet. An der Metalloberfläche können auch neutrale Moleküle adsorbiert werden. Das betrifft bevorzugt Dipolmoleküle, z. B. des Wassers, die sich dabei an der Phasengrenze gerichtet anlagern. Die adsorbierten Moleküle können sich zwischen das Metall und die solvatisierten Überschußionen schieben und verlagern damit die äußere Helmholtzfläche nach dem Inneren der Lösung (Abb. 7.3.4-9).

Die Dipole leisten ebenfalls einen Beitrag zur Größe des Potentialsprungs an der Phasengrenzfläche. Daher setzt sich die Galvanispannung additiv aus den beiden durch Überschußladungen und durch Dipole verursachten An-

teilen zusammen. Die besondere Struktur der Dipolschicht beeinflußt in starkem Maße die Ladungsverteilung innerhalb der Doppelschicht. Gerichtete Dipolschichten weisen gegenüber Schichten mit unregelmäßiger Struktur ein vollständig anderes Verhalten, z. B. hinsichtlich ihrer Dielektrizitätskonstanten, auf. Für Wasserschichten mit gut ausgerichteten Dipolen kann man beispielsweise bei Zimmertemperatur eine Dielektrizitätskonstante von etwa 6 annehmen, während sie für Wasser bei regelloser Verteilung der Moleküle rund 80 beträgt.

Das Modell der elektrochemischen Doppelschicht erinnert stark an zwei in Reihe geschaltete Kondensatoren, so daß für die *Doppelschichtkapazität* C_{ges} gilt $1/C_{ges} = 1/C_{starr} + 1/C_{diffus}$. Die Kapazität C_{starr} der starren Schicht hängt maßgeblich von ihrer Dicke ab, die durch die Abmessungen der spezifisch adsorbierten, teilweise desolvatisierten Ionen und der solvatisierten Überschußionen beeinflußt wird.

Noch komplizierter ist der Aufbau der Doppelschicht an Grenzflächen zwischen Halbleiter und Elektrolytlösung. Dort kann auch der im Halbleiter liegende Teil der Doppelschicht einen diffusen Anteil haben, der durch die im Vergleich zu Metallen geringere Ladungsträgerkonzentration zustande kommt. Solche diffusen Zonen haben Einfluß auf das elektrochemische Verhalten von Halbleitern. Der Ladungsverlauf ist ähnlich dem an einer Halbleiter-Metall-Grenzfläche, wie er sich nach einer Kontaktierung herausbildet (↑ Abb. 8.3.6-7).

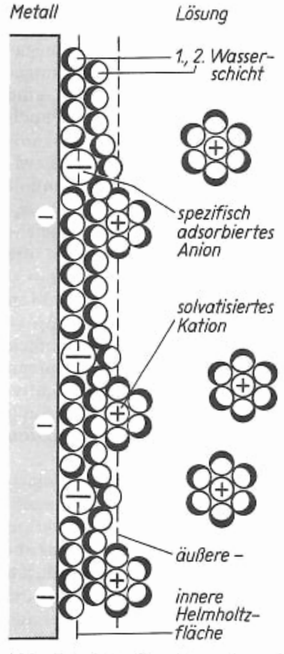


Abb. 7.3.4-9 Struktur der elektrochemischen Doppelschicht (Modell nach BOCKRIS, DEVANATHAN und MÜLLER), die positive Seite der Dipolmoleküle ist durch *Randverdickung* kenntlich gemacht

Die Doppelschichtkapazität läßt sich durch Wechselstrom- oder Impulsmethoden messen, wie sie auch für normale Kondensatoren angewandt werden. Im Unterschied zu letzterem ist sie jedoch von der Ladungsdichte und damit von der Galvanispannung abhängig. Sie hat die Größenordnung von 5 bis 50 $\mu F cm^{-2}$. Der Verlauf der Doppelschichtkapazität als Funktion der Galvanispannung erlaubt Aussagen über die Struktur der Doppelschicht (↑ Abb. 7.3.4-8, 7.3.4-9).

7.3.4.2. Zeitlicher Verlauf elektrochemischer Reaktionen. Beim Ablauf einer elektrochemischen Reaktion treten stets Hemmungen auf, die so stark sein können, daß trotz großer Entfernung vom Gleichgewichtszustand keine merklichen Veränderungen in dem System stattfinden. Solche Hemmungen können erwünscht sein, z. B. bei elektrochemischen Korrosionsvorgängen, oder aber unerwünscht, weil zu ihrer Überwindung, z. B. bei technischen Elektrolysen, ein großer Energieaufwand nötig ist. Für eine gezielte Beeinflussung elektrochemischer Reaktionen in der einen oder anderen Richtung ist eine genaue Kenntnis der Teilvorgänge und der Art ihrer Hemmungen erforderlich. Das daraus resultierende große wissenschaftliche und technische Interesse hat zu einer schnellen Entwicklung der elektrochemischen Kinetik in den letzten Jahrzehnten geführt.

7.3.4.2.1. *Reaktionsstromstärke* Die Geschwindigkeiten elektrochemischer Reaktionen können wie die chemischer Reaktionen durch den Quotienten aus der Objektmenge der Formelumsätze $d\xi$ und dem zugehörigen Zeitintervall dt beschrieben werden (↑ 7.2.) Da der Stoffumsatz bei elektrochemischen Reaktionen zugleich mit einem Ladungsdurchtritt durch die Phasengrenze der jeweiligen Elektrode verbunden ist, entspricht die Reaktionsgeschwindigkeit einer Stromstärke $I = dq/dt$, mit der Elektrizitätsmenge q. Unter Berücksichtigung der Faradayschen Gesetze gilt der Zusammenhang $I = z_r F\, d\xi/dt$. $z_r = dq/(F\, d\xi)$ ist die Reaktionsladungszahl, I die *Reaktionsstromstärke*, die leicht gemessen werden kann. Der Reaktionsstrom oder *Faradaysche Strom* führt zu einem chemischen Umsatz an den Elektroden, während *nichtfaradaysche Ströme* z. B. zur Auf- und Entladung der Doppelschichtkapazität fließen. Im stationären Zustand einer gleichstromdurchflossenen Elektrode ist die Gesamtstromstärke stets gleich der Reaktionsstromstärke. Man bezieht den Ladungsdurchtritt oft auf die Einheit der Grenzfläche A einer Elektrode und verwendet

als aussagekräftiges Maß die *Stromdichte* $i = I/A$.
Je nach Richtung des Ladungstransports unterscheidet man zwischen *anodischen* (positiven) und *katodischen* (negativen) *Strömen*. Ein anodischer Strom ($I > 0$) fließt, wenn positive Ladungsträger von der Metallphase einer Elektrode, der *Anode*, in die Elektrolytlösung oder

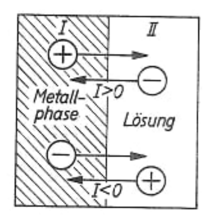


Abb. 7.3.4-10 Festlegung der Stromrichtung

negative Ladungsträger in umgekehrter Richtung übergehen. Ein katodischer Strom ($I < 0$) fließt bei umgekehrter Durchtrittsrichtung der Ladungsträger. Diese Elektrode heißt dann *Katode*. An einer Anode findet stets eine Oxydation, an einer Katode eine Reduktion statt (Abb. 7.3.4-10).
Der gesamte Stoffumsatz an einer Elektrode setzt sich aus den anodischen und den katodischen Teilumsätzen zusammen. An einer Zinkelektrode $Zn(I)/Zn^{2+} \cdot aq(II)$ läuft z. B. die anodische Teilreaktion $Zn^{2+}(I) \rightarrow Zn^{2+} \cdot aq(II)$ mit der Teilstromstärke $\vec{I} = 2F\,d\xi/dt$ ab. Für die katodische Teilreaktion $Zn^{2+} \cdot aq(II) \rightarrow Zn^{2+}(I)$ ist die Teilstromstärke negativ: $\overleftarrow{I} = -2F\,d\xi/dt$. Die Summe beider ergibt die Gesamtstromstärke $I = \vec{I} + \overleftarrow{I}$ oder bezogen auf die Fläche die Gesamtstromdichte $i = \vec{i} + \overleftarrow{i}$. Speziell im Gleichgewicht ist $i = 0$, d. h., die Stromdichten von Hin- und Rückreaktion sind gleich groß; ihre Beträge werden dann als *Austauschstromdichte* i_0 bezeichnet: $i_0 = |\vec{i}| = |\overleftarrow{i}|$.
Die Austauschstromdichten charakterisieren die elektrochemische Reaktivität. Hohe Austauschstromdichten in der Größenordnung von Acm^{-2} zeichnen katalytisch aktive Elektroden aus, die bei Elektrolysezellen und Brennstoffelementen angestrebt werden. Auch sehr kleine Austauschstromdichten von 10^{-12} Acm^{-2} wurden beobachtet. An *einfachen Elektroden* mit so kleinen Austauschstromdichten stellt sich das Gleichgewicht nach einer Verschiebung nur sehr langsam wieder ein. An *mehrfachen Elektroden*, an denen mehrere elektrochemische Reaktionen nebeneinander ablaufen, entfallen auf die einzelnen Reaktionen bestimmte Teilstromstärken. Eine mehrfache Elektrode ist z. B. die Zinkelektrode in saurer Lösung, wenn neben der Auflösung von Zink eine Abscheidung von Wasserstoff stattfindet: $Zn(I) \rightarrow Zn^{2+}(II) + 2e^-(I)$ mit der Teilstromstärke I_{Zn} und $2\,H^+(II) + 2e^-(I) \rightarrow H_2(II)$ mit der Teilstromstärke I_H. Auch hier ist der Gesamtstrom wieder die Summe der Teilströme: $I = I_{Zn} + I_H$. Taucht man ein Zinkblech einfach in eine verdünnte Säurelösung, kann insgesamt kein Strom durch die Elektrode fließen und die Beträge der Teilstromstärken müssen gleich sein $I_{Zn} = |I_H|$. Der Elektronenaustausch zwischen den Reaktionsteilnehmern erfolgt dann über das Elektrodenmetall. Bei technischen Prozessen und in der Analytik ist gewöhnlich nur eine einzige von mehreren parallelen Elektrodenreaktionen von Interesse. Der Anteil dieser Hauptreaktion an der Gesamtstromstärke I wird durch die *Stromausbeute* η charakterisiert $\eta = I_{Haupt}/I$.

7.3.4.2.2. *Überspannung und Polarisation.* Eine elektrochemische Reaktion kann nur dann im Gleichgewicht sein, wenn insgesamt kein Stoff- und Energieumsatz mehr stattfindet, wenn also kein Strom die Elektrode durchfließt. Ist das nicht der Fall, so treten positive oder negative Abweichungen von der Gleichgewichtsgalvanispannung auf. Wird z. B. die *einfache* Zinkionenelektrode $Zn(I)Zn^{2+} \cdot aq(II)$ in einer galvanischen Zelle mit dem negativen Pol einer Spannungsquelle verbunden, so findet die Reaktion $Zn^{2+} + 2e^- \rightarrow Zn$ statt, die Elektrode wird von einem negativen Strom durchflossen. Dazu ist eine Galvanispannung erforderlich, die negativer als der Gleichgewichtswert ist. Bei umgekehrter Polung der Zinkelektrode findet eine Auflösung des Zinks statt, der Strom und auch die Abweichung von der Gleichgewichtsgalvanispannung sind dann positiv. Die Differenz zwischen der Galvanispannung g_I bei Stromfluß und der Gleichgewichtsgalvanispannung g_{gl} wird als Überspannung $\eta = g_I - g_{gl}$ bezeichnet. Bei *mehrfachen Elektroden* entspricht die Galvanispannung g_0 bei der Stromstärke $I = 0$ in der Regel keinem Gleichgewichtswert. Man bezeichnet dann allgemein die Differenz $\Delta g = g_I - g_0$ als *Elektrodenpolarisation* oder einfach als *Polarisation*. Sie ist bei anodischen Strömen stets positiv, bei katodischen Strömen negativ. Im allgemeinen Sprachgebrauch wird häufig nicht zwischen Überspannung und Polarisation unterschieden.
Die Ursache der Überspannung und der Polarisation sind Hemmungen von Teilvorgängen, aus denen sich jede elektrochemische Reaktion zusammensetzt. Bei der katodischen Zinkabscheidung müssen die Zinkionen zunächst durch Diffusion aus dem Inneren der Lösung an die Doppelschicht gelangen, den diffusen Bereich der Doppelschicht passieren und in die äußere Helmholtzfläche eingereiht werden, ehe der Ladungsdurchtritt erfolgen und das Zinkion in das Elektrodenmetall übertreten kann. Besteht dieses aus flüssigem Zinkamalgam, so schließt sich dort noch eine Diffusion in das Innere des

Amalgams an, ist es reines Zink, muß die Einordnung in das Kristallgitter erfolgen. Je nachdem, welcher der angeführten Vorgänge besonders stark gehemmt wird und damit die Ursache einer Überspannung ist, unterscheidet man *Diffusions-*, *Durchtritts-* und *Kristallisationsüberspannung*. Wenn der Ladungsdurchtrittsreaktion noch eine rein chemische Reaktion vor- oder nachgelagert ist, kann auch eine *Reaktionsüberspannung* auftreten. Eine solche Reaktion kann z. B. die Solvatation eines Ions, eine Komplexbildung oder ein Dissoziationsvorgang sein. Die Größe der Überspannung oder Polarisation hängt von der *Stromdichte* ab. Entsprechende Zusammenhänge werden gewöhnlich aus zeitlich konstanten, d. h. *stationären Werten* von Stromdichte und Überspannung gewonnen. Aus ihnen sind häufig Schlüsse auf die am stärksten gehemmte Teilreaktion zu ziehen. Weiterreichende Aufschlüsse über die Reaktionsmechanismen liefern in der Regel *nichtstationäre Methoden*, bei denen eine Variable, z. B. die Stromdichte i oder die Überspannung η, periodisch oder aperiodisch verändert und die Antwortfunktion von η bzw. i verfolgt wird. Diese Messungen können mit einem großen apparativen Aufwand verbunden sein, wenn entweder sehr kleine Änderungen oder sehr schnell ablaufende Vorgänge erfaßt werden müssen.

Elektrochemische Reaktionen unterscheiden sich von rein chemischen heterogenen Reaktionen in einem charakteristischen Teilschritt, nämlich dem Durchtritt eines Ladungsträgers durch die Phasengrenzfläche. Diese *Durchtrittsreaktion* ist bei einer Redoxreaktion, z. B. bei der Reduktion von Eisen(3)-Ionen nach $Fe^{3+}(II) + e^-(I) \rightarrow Fe^{2+}(II)$, der Übergang des Elektrons aus der metallischen Phase (I) auf das Eisen(3)-Ion, das sich in der inneren oder äußeren Helmholtzschicht befindet.

Bei einer Zinkamalgamelektrode mit der Reaktion $Zn^{2+} \cdot aq(II) + Hg(I) + 2e^-(I) \rightarrow Zn(Hg)\ (I)$ (Abb. 7.3.4-11) muß das Zinkion vor dem Übergang in das Elektrodenmetall zunächst dehydratisiert werden. Gewöhnlich wird das Ion adsorbiert und verliert dabei einen Teil seiner Hydrathülle. Man spricht von einem *Ad-Ion*, wenn es seine Ladung noch besitzt, von einem *Ad-Atom*, wenn es entladen ist. Meist bildet sich ein Zwischenzustand heraus, bei dem das Ion formal einen Teil seiner Ladung abgegeben hat. Man spricht dann von einem *partiellen Ladungsübergang*. Der letzte Schritt ist der Übertritt in das flüssige Elektrodenmetall, in dem es sich von den übrigen Metallatomen der Grenzfläche nicht mehr unterscheidet.

Die Durchtrittsgeschwindigkeit der Ladungsträger durch die Phasengrenzfläche hängt von dem Gefälle des elektrischen Potentials über der Durchtrittsschicht und damit von der Galvanispannung der Elektrode ab. Bei der Gleichgewichtsgalvanispannung treten z. B. bei der Zinkionenelektrode in der Zeiteinheit genausoviel Zinkionen aus der Lösung in das Metall über, wie sich aus dem Metall ablösen und in die Lösungsseite der Doppelschicht einreihen. Verändert man die Galvanispannung, indem man der Elektrode eine positive Überspannung erteilt, so werden die Zinkionen schneller in Lösung gehen, die positive, anodische Stromrichtung wird dadurch begünstigt. Quantitative Aussagen lassen sich machen, wenn man berücksichtigt, daß nach der *Theorie des Übergangszustands* (↑ 7.2.2.) jedes Teilchen X_i bei der Durchtrittsreaktion einen Energieberg über-

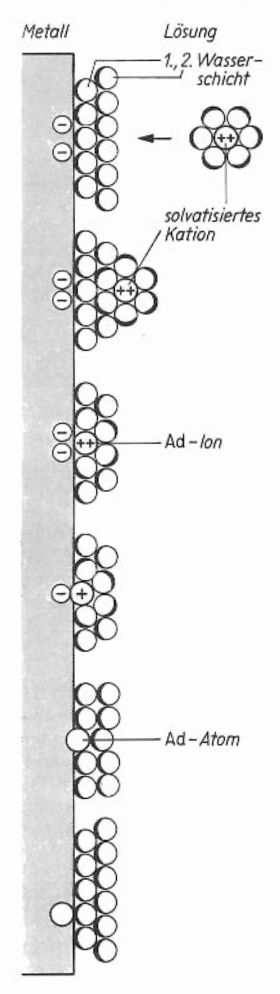


Abb. 7.3.4-11 Katodische Abscheidung eines Metalls an Quecksilber

winden muß. Die Höhe des Energiebergs, gemessen von dem jeweiligen Ausgangszustand des Ladungsträgers, gibt die *Aktivierungsenergie* $\vec{E}_{gl}$ bzw. $\overleftarrow{E}_{gl}$ der Hin- bzw. der Rückreaktion im Gleichgewicht an, die wiederum die Austauschstromdichte i_0 bestimmt (Abb. 7.3.4-12, durchgezogene Kurve). Wird die Galvanispannung um den Betrag der *Durchtrittsüberspannung* η_D variiert, so muß sich die freie elektrochemische Standardenthalpie der Durchtrittsreaktion $\Delta\tilde{G}^{\ominus}_{gl}$ um den Betrag $z_iF\eta_D$ verändern. Dabei wird der Energieberg verformt (Abb. 7.3.4-12, gestrichelte Kurve) und die Aktivierungsenergie der

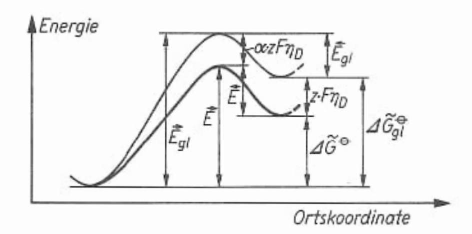


Abb. 7.3.4-12 Schematische Darstellung des Energiebergs bei Durchtrittsreaktionen; *dünne Kurve* ohne, *dicke Kurve* mit Überspannung; α Durchtrittsfaktor $0 < \alpha < 1$

anodischen Reaktion $\vec{E}_{gl}$ verringert sich um den Betrag $\alpha z_iF\eta_D$ nach $\vec{E}$, die Aktivierungsenergie der katodischen Reaktion $\overleftarrow{E}_{gl}$ erhöht sich um den Betrag $(1 - \alpha)\, z_iF\eta_D$ auf den Wert $\overleftarrow{E}$. α wird als *Durchtrittsfaktor* bezeichnet und hat die Größe $0 < \alpha < 1$. Der Durchtrittsfaktor kann in einem gewissen Bereich der Galvanispannung als konstant angesehen werden. Zwischen der Stromdichte i bei der Überspannung η_D und der Austauschstromdichte i_0 im Gleichgewicht besteht der Zusammenhang:
$i/i_0 = (z_i/|z_i|)\,\{\exp\,[\alpha z_iF\eta_D/(RT)] - \exp\,[-(1-\alpha)\, z_iF\eta_D/(RT)]\}$.
Bei der Ableitung dieser Beziehung wurde vorausgesetzt, daß nur Durchtrittsüberspannung auftritt. Entwickelt man die Exponentialfunktion in eine Reihe, so resultiert für sehr kleine Durchtrittsüberspannungen $i = i_0\,|z_i|\,F\eta_D/(RT)$; d. h. ein linearer Zusammenhang zwischen der Stromdichte i und der Durchtrittsüberspannung η_D.
Für große Überspannungen $\eta_D \gg RT/(|z_i|F)$ läßt sich jeweils eines der Exponentialglieder vernachlässigen, und es folgt $\ln i = \ln i_0 + (\alpha\,|z_i|\,F\eta_D)/(RT)$ für $i > 0$ oder $\ln |i| = \ln i_0 - (1-\alpha)\,|z_i|\,F\eta_D/(RT)$ für $i < 0$. Diese Beziehungen der Form $\eta_D = a + b \ln |i|$ werden als *Tafelgleichungen* bezeichnet. Aus den Konstanten a und b lassen sich durch Koeffizientenvergleich α und i_0 ermitteln; α liegt häufig bei 0,5.
Wenn bei einer elektrochemischen Reaktion der eigentliche Ladungsübergang wenig gehemmt verläuft, so tritt in unmittelbarer Nähe der elektrochemischen Doppelschicht häufig eine Verarmung bzw. Anreicherung der Reaktanten ein. Die Konzentrationsunterschiede gegenüber dem Inneren der Phase versuchen sich wegen des Gefälles des elektrochemischen Potentials auszugleichen. Das elektrochemische Potentialgefälle $d\tilde{\mu}_i/dx = d\mu_i/dx + z_iF\,d\varphi/dx$ einer Teilchensorte X_i in der Richtung x senkrecht zur Elektrodengrenzfläche hat zwei Anteile. Der erste bewirkt eine *Diffusion*, d. h. eine Teilchenwanderung im Gefälle des chemischen Potentials, der zweite eine *Migration* oder *Überführung*, d. h. eine Wanderung im Gefälle des elektrischen Potentials. Diesen Transportvorgängen entspricht die Stromdichte $-i_i = z_iFD\,dc_i/dx + c_i|z_i|\Lambda_i\,d\varphi/dx$. Der erste Summand auf der rechten Seite enthält den Diffusions-, der zweite den Migrationsanteil. $-d\varphi/dx$ ist die elektrische Feldstärke, dc_i/dx das *Konzentrationsgefälle* der Teilchensorte X_i und D_i ihr Diffusionskoeffizient. Werden *Moleküle* an der Elektrode umgesetzt, oder ist ein inertes *Leitsalz* in großem Überschuß vorhanden, kann der Anteil der Migration vernachlässigt werden. Dann bildet sich eine Diffusionsschicht der Dicke δ aus, der man nach NERNST ein konstantes Konzentrationsgefälle $dc_i/dx = (c_i^{\mathrm{II}} - c_i^0)/\delta$ zuschreibt; die Konzentration c_i^0 ist am Rande der Doppelschicht, c_i^{II} im Inneren der Lösung vorhanden. Die maximal mögliche Stromdichte wird dann erreicht, wenn c_i^0 auf den Wert Null absinkt. Man spricht dann von der *Diffusionsgrenzstromdichte* i_d^{gr}, die konzentrationsproportional ist und deshalb für analytische Untersuchungen, z. B. bei der *Polarographie*, genutzt wird. Eine typische polarographische Strom-Spannungs-Kurve mit einer Quecksilbertropfelektrode als Katode zeigt Abb. 7.3.4-13. Wegen der Form dieser Kurve spricht man von einer *polarographischen Stufe*. Bei Konzentrationsbestimmungen aus Diffusionsgrenzströmen muß man für definierte Konvektionsbedingungen sorgen. Erhöhung der Konvektion, z. B. durch verstärkte Rührung, führt zu einer Verringerung der Diffusionsschichtdicke und damit zu einer Erhöhung des Grenzstroms. Die *Diffusionsüberspannung* $\eta_d = g_1 - g_{gl}$ wird aus der Nernstschen Gleichung zugänglich, wenn man zur Berechnung von g_1 die gegenüber dem stromlosen Zustand veränderte Konzentration c_i^0 am Rande der Doppelschicht in die Gleichung einsetzt $\eta_d = [RT/(z_iF)]\ln(c_i^0/c_i^{\mathrm{II}}) = RT/(z_iF)\ln(1 - i/i_d^{\mathrm{gr}})$.
Häufig sind der Durchtrittsreaktion neben den Transportvorgängen auch homogene oder heterogene chemische Reaktionen vorgeschaltet oder nachgelagert. Solche Reaktionen führen zu

einer *Reaktionsüberspannung* η_R, wenn sie die gehemmtesten Teilreaktionen darstellen. Dabei treten infolge der endlichen Reaktionsgeschwindigkeit Abweichungen vom Gleichgewichtszustand in Form von Konzentrationsänderungen im Bereich der *Reaktionsschicht* auf. Wegen dieser Ähnlichkeit mit der Diffusionsüberspannung faßt man Diffusions- und Reaktionsüberspannung unter dem Oberbegriff *Konzentrationsüberspannung* zusammen.

Heterogene chemische Teilreaktionen können auch unmittelbar an der Elektrodenoberfläche, z. B. unter Einbeziehung adsorbierter Teilchen, ablaufen. Dann liegen die Konzentrationsabweichungen von dem Gleichgewichtswert und damit die Reaktionszone im Bereich der elektrochemischen Doppelschicht. Bei der Wasserstoffentwicklung an Metallelektroden kann sich dem Durchtrittsvorgang, der *Volmerreaktion* $H_3O^+ + e^- + M \rightleftharpoons MH_{ads} + H_2O$, die Rekombination zweier an dem Elektrodenmetall M adsorbierter Wasserstoffatome MH_{ads} in der *Tafelreaktion* anschließen: $2\,MH_{ads} \rightleftharpoons 2\,M + H_2$. Ist die Tafelreaktion gehemmt, so ist die Konzentration von MH_{ads} stromdichteabhängig und größer als im Gleichgewicht. Das führt zu einer Reaktionsüberspannung, die sich durch eine *Tafelbeziehung* quantitativ ausdrücken läßt: $\eta_R = a - (RT/2F) \ln |i_R|$. Der Summand a enthält katalytische Eigenschaften des Elektrodenmetalls. Er hängt auch von dessen Oberflächenzustand und der Temperatur ab. Für heterogene vor- oder nachgelagerte Reaktionen werden auch Grenzströme beobachtet. Zum Unterschied von Diffusionsgrenzströmen sind *Reaktionsgrenzströme* nicht rührabhängig.

Bei der katodischen Abscheidung von Metallen an festen Elektroden folgt der Durchtrittsreaktion noch der Einbau des Ad-Atoms in das Metallgitter. Nach der Theorie von KOSSEL und von STRANSKI gibt es verschiedene Lagen der

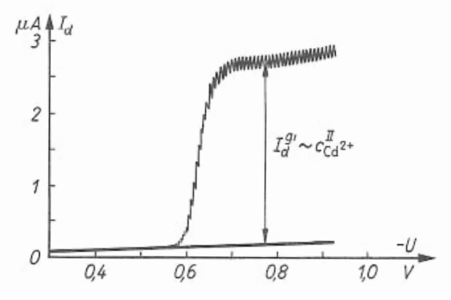


Abb. 7.3.4-13 Polarogramm einer Cadmiumionenlösung in 0,1 n Kaliumchlorid als Leitelektrolyt; *dicke Kurve* Grundstrom in reiner Leitelektrolytlösung, *dünne Kurve* polarographische Stufe von Cadmium(2)-Ionen, I_d^{gr} Maß für die Cd^{2+}-Konzentration in der Lösung

Atome auf Kristallen; Ad-Zustand *5*, Stufenlage *4*, Wachstumsstelle *1* (Abb. 7.3.4-14). Die Durchtrittsreaktion erfordert in der Lage *5* die geringste Aktivierungsenergie, weil hier die Solvatation weitgehend erhalten bleibt und hat so an dieser Stelle die größte Wahrscheinlichkeit. Der Einbau ins Metallgitter erfolgt meistens an der Wachstumsstelle *1* nach Oberflächendiffusion und Desolvatisierung des Ad-Atoms, weil dort die größte Anlagerungsenergie frei wird. Sind Oberflächendiffusion oder Einbau in das Metallgitter die am stärksten gehemmten Vorgänge bei der katodischen Metall-

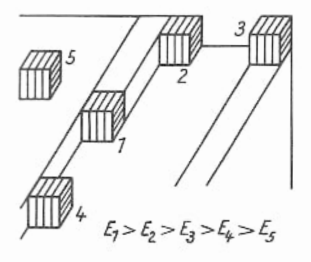


Abb. 7.3.4-14 Relativer Energiegewinn bei der Anlagerung von Ionen an verschiedenen Stellen einer Kristalloberfläche

abscheidung, so tritt eine *Kristallisationsüberspannung* auf. Experimentell findet man für sehr kleine Kristallisationsüberspannungen eine lineare Abhängigkeit von der Stromdichte. Im mittleren Bereich gilt gewöhnlich eine Tafelgleichung. Die Abscheidung fester Metalle, die in der Galvanotechnik und Metallurgie vorgenommen wird, bezeichnet man als *Elektrokristallisation*. Dabei ist eine Beeinflussung der Abscheidung durch *Kristallisationsinhibitoren* möglich, die sich an Wachstumsstellen anlagern, damit die Kristallisationsüberspannung erhöhen und eine Abscheidung des Metalls über die Bildung neuer Kristallkeime begünstigen. Dies führt zur Entstehung feinkristalliner, glatter oder glänzender Metallniederschläge.

Auch bei anodischen Auflösungsvorgängen können Kristallisationsüberspannungen auftreten. Der Ausbau aus dem Gitter erfolgt nicht ungehemmt, sondern erfordert eine Aktivierungsenergie, die der Differenz der Energien zwischen eingebautem und Ad-Atom entspricht. Dieser Teilschritt ist unabhängig von der Überspannung, so daß ähnlich wie bei der Konzentrationsüberspannung Grenzströme auftreten.

Neben der Überspannung kann zusätzlich eine *Widerstandspolarisation* auftreten, d. h. ein Potentialabfall über schlecht leitende Elektrodengrenzschichten. Bei Elektrolytlösungen und -schmelzen hängt die Widerstandspolarisation von den Äquivalentleitfähigkeiten Λ_i und den

Konzentrationen c_i der an dem Ladungstransport beteiligten Teilchen ab und hat entsprechend dem Ohmschen Gesetz die Größe $\Delta\varphi = is/\varkappa = is/\sum(\Lambda_i c_i |z_i|)$; s Schichtdicke. In festen Deckschichten auf Elektroden kann Elektronen- oder Ionenleitfähigkeit vorliegen. Bei Elektronenleitfähigkeit gilt ebenfalls das Ohmsche Gesetz, $\Delta\varphi = is\varrho$, wobei ϱ der spezifische Widerstand der festen Phase ist. Für ionenleitende Deckschichten ist die Leitfähigkeit feldstärkeabhängig und folgt oftmals der Beziehung $\Delta\varphi = sk(\ln i - \ln i_0)$; k und i_0 sind materialabhängige Konstanten. Widerstandspolarisation ist stets in der Gesamtpolarisation einer Elektrode enthalten.

7.3.4.3. Anwendung elektrochemischer Prozesse. Elektrochemische Prozesse finden in Wissenschaft und Technik breite Anwendung und haben große Bedeutung für Lebensvorgänge. Ihre Vielseitigkeit, aber auch Kompliziertheit, spiegelt sich in der stürmischen Entwicklung der Wissenschaftsdisziplin Elektrochemie in den vergangenen Jahrzehnten wider. Thermodynamische und kinetische Gesetzmäßigkeiten elektrochemischer Reaktionen werden weit über den Bereich der Elektrochemie hinaus genutzt. In der Produktion spielen Elektrolysen, z. B. zur Herstellung von Chlor und Alkalilaugen, zur Metallgewinnung und -raffination, eine große Rolle. Die Verwendung tragbarer elektronischer Geräte ist eng mit der Herstellung galvanischer Elemente als Energielieferanten verbunden, und die Einsatzbereitschaft von Kraftfahrzeugen hängt weitgehend von der Leistungsfähigkeit elektrochemischer Stromquellen ab.

Elektrochemische Vorgänge haben besondere Bedeutung für die Aufklärung und Vermeidung von Korrosionsvorgängen an Metallen. Millionenverluste durch Metallkorrosion werden auf der Welt durch die Anwendung anodischer und katodischer Verfahren des Korrosionsschutzes oder durch galvanisch hergestellte, korrosionshemmende Überzüge verhindert. Für die Herstellung kompliziert geformter metallischer Werkstücke benutzt man die elektrochemische Metallbearbeitung, bei der durch anodische Auflösung, oder die Galvanoformung, bei der durch katodische Abscheidung, eine Formgebung erreicht wird. In der Analytik, der Meß-, Steuer- und Regelungstechnik bedient man sich gern elektrochemischer Verfahren, wobei sicher die Messung und Regelung von pH-Werten die bekannteste Anwendung darstellt. Viele Stoffwechselvorgänge in lebenden Organismen werden durch Donnanspannungen an Membranen von Zellen geregelt, und die Weiterleitung von Reizen in Nerven erfolgt durch elektrochemische Mechanismen.

8. Festkörper

8.1. Gitterdynamik der Festkörper

8.1.1. Gegenstand der Festkörperphysik

Feste Stoffe besitzen aufgrund ihrer Formbeständigkeit eine überragende Bedeutung als *Werkstoffe* in den verschiedenen Gebieten der Technik. Bis weit in unser Jahrhundert hinein gründeten sich die Herstellung, Verarbeitung und Anwendung der Werkstoffe im wesentlichen auf praktische Erfahrungen und die phänomenologische Wiedergabe von Zusammenhängen zwischen Stoffeigenschaften und Zustandsgrößen. Erst auf der Grundlage der Atomphysik und der Quantenmechanik wurde es prinzipiell möglich, Struktur und Eigenschaften der Festkörper mit Hilfe von Modellen über die Anordnung der atomaren Bausteine und den zwischen ihnen wirkenden Kräften im eigentlichen Sinne zu deuten. Damit bildete sich die *Festkörperphysik* als umfangreiche und praktisch äußerst bedeutungsvolle Teildisziplin der Physik heraus.

Ziel und Gegenstand der Festkörperphysik ist es, die außerordentliche Vielfalt der Eigenschaften und Umwandlungen fester Stoffe mittels atomistischer Modelle qualitativ und auch quantitativ zu interpretieren und umgekehrt, Voraussagen über neue, meist für technische Anwendungen interessante Effekte abzuleiten. Ergänzt wird die Festkörperphysik durch die *Festkörperchemie*, die sich mit chemischen Vorgängen im festen Zustand befaßt.

Makroskopische Festkörper bestehen aus einer sehr großen Anzahl von atomaren Bausteinen, die überdies durch intensive und teilweise gerichtete Wechselwirkungskräfte miteinander verbunden sind. Um Festkörpereigenschaften atomistisch zu beschreiben, muß man daher die Beiträge vieler Teilchen summieren. Ebenso wie bei der theoretischen Behandlung anderer *Teilchensysteme* (↑ 6.1.) leisten dabei die mathematischen Methoden der *statistischen Physik* wertvolle Dienste. Als gebundene atomare Systeme unterliegen die dynamischen Erscheinungen im Festkörper den Gesetzmäßigkeiten der Quantentheorie (s. a. 2.5.1.). Kennzeichnend für die Festkörperphysik ist es, daß *Quanteneffekte* wegen der starken Kopplung vieler Teilchen auch über größere Dimensionen in Erscheinung treten. Das gilt besonders für die atomaren Schwingungen und die dadurch bedingten thermischen Festkörpereigenschaften sowie für die Besetzung der Elektronenzustände im Festkörper und die daraus resultierenden elektrischen und optischen Prozesse.

Besonders übersichtliche Verhältnisse liegen in Festkörpern vor, bei denen die atomaren Bausteine streng in der translationssymmetrischen Struktur eines *Kristallgitters* (↑ 6.4.1.) angeordnet sind. Ein wichtiges Ausgangsmodell für die theoretische Beschreibung bildet daher der *ideale Festkörper*, d. h. ein unendlich ausgedehnter *Einkristall*, der weder Kristalldefekte noch Verunreinigungen enthält und in dem die Atome keine thermische Bewegung ausführen. Läßt man die letzte Einschränkung fallen und berücksichtigt thermische Schwingungen der Atome um ihre Gleichgewichtslagen, so ergibt sich das schon realistischere Modell des *perfekten Festkörpers*. Nach dem *Blochschen Theorem* müssen alle Funktionen, die mikrophysikalische Eigenschaften im Kristallgitter beschreiben, wie beispielsweise der Verlauf des Gitterpotentials oder die Wellenfunktionen der Kristallelektronen, im idealen Festkörper die gleiche dreifache Periodizität aufweisen wie die Anordnung der atomaren Bausteine, die durch *Translationssymmetrie* (s. a. 6.4.8.) festgelegt ist:

$$(1) \quad \boldsymbol{T} = n_1\boldsymbol{a} + n_2\boldsymbol{b} - n_3\boldsymbol{c}.$$

Dabei sind n_1, n_2, n_3 ganze Zahlen und $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ bezeichnen die Gittergrundvektoren. Praktisch läßt sich das Modell des idealen Festkörpers am ehesten durch defektarme große Einkristalle hoher Reinheit annähern; derartige Proben, die von den Gegebenheiten natürlicher fester Stoffe relativ weit entfernt sind, bilden daher bevorzugte Objekte der festkörperphysikalischen Grundlagenforschung. Im Ergebnis haben Einkristalle hoher Perfektion und Reinheit, aber auch als hochveredelte Werkstoffe der Optik und Mikroelektronik wichtige Anwendungen gefunden.

Ausgehend von dem zunehmenden Verständnis des Verhaltens einkristalliner Proben, hat sich die Festkörperphysik mehr und mehr der Erforschung des Einflusses von Kristalldefekten zugewandt. Dabei zeigte es sich, daß die praktisch bedeutsamen Werkstoffeigenschaften häufig sehr stark, manchmal sogar vorherrschend, von den Kristalldefekten oder der *Realstruktur* (s. a. 6.4.7.) des Festkörpers bestimmt werden. In dem Maße, wie man es lernte, auch weniger geordnete Festkörper strukturell zu erforschen und ihre Eigenschaften gleichfalls atomistisch

zu beschreiben, hat sich der Gegenstand der Festkörperphysik erheblich erweitert. Damit wurde eine noch andauernde Entwicklung eingeleitet, in deren Ergebnis erwartet werden darf, daß die in wichtigen Teilgebieten der Werkstofftechnik noch vorherrschende empirisch-qualitative Beschreibung durch die strenge festkörperphysikalische Untermauerung vervollständigt wird. Teilgebiete, in denen das fruchtbare Zusammenwirken von Werkstoffwissenschaften und Festkörperphysik zunehmend in neuen technischen Lösungen zum Tragen kommt, bilden Gläser und Baustoffe, organische Kristalle und Plaste und vor allem die metallischen Legierungen sowie Verbundwerkstoffe zwischen Metallen und anderen Materialien. Weiterhin bildet die Erforschung und Nutzung von Erscheinungen, die ganz oder vorwiegend an *Oberflächen* oder in *dünnen Schichten* der festen Stoffe ablaufen, einen Schwerpunkt der gegenwärtigen Festkörperphysik. Als Nahtstellen zwischen Werkstoff und Umwelt spielen die Oberflächen eine wichtige Rolle, und ihre gezielte Verbesserung durch *Implantieren* von Fremdstoffen oder durch *Beschichten* mit widerstandsfähigen Materialien verspricht einen wesentlichen Beitrag bei der Erhöhung der Härte und Korrosionsbeständigkeit von metallischen Erzeugnissen und damit zur Lösung von Problemen der *Materialökonomie* bei Einsparung knapper Bestandteile im Werkstoffvolumen. *Dünne Schichten* ermöglichen in der Optik und Elektronik die Realisierung wichtiger Prozeßschritte der Verarbeitung und Speicherung von Informationen bei sparsamstem Materialeinsatz.

Sowohl für Volumenwerkstoffe als auch als dünne oder dicke Schichten haben in den letzten Jahren die *amorphen* oder *teilamorphen* Phasen wachsendes Interesse erlangt, die von bestimmten Metallegierungen oder von Halbleitern wie Silicium durch eine spezielle Abscheidungstechnik bei sehr raschem Abkühlen gewonnen werden können. Die Skala verfügbarer Werkstoffe wird durch diese *metastabilen Phasen* in wünschenswerter Weise bereichert, und es zeichnen sich viele Anwendungsmöglichkeiten ab, die von neuartigen Lösungen für magnetische Abschirmungen bis zu korrosionsbeständigen Überzügen, von großflächigen Siliciumsolarzellen für Sonnenkraftwerke bis zu Hartstoffschichten mit diamantähnlichen Eigenschaften reichen.

Nach einer Prognose des Nobelpreisträgers ANDERSON wird das gegenwärtige Jahrzehnt in der Entwicklung der Festkörperphysik durch die zunehmende Erforschung und Nutzung von Stoffzuständen mit hoher atomarer Unordnung gekennzeichnet sein. Daneben zeichnen sich freilich auch Tendenzen ab zur Schaffung komplizierter „künstlich geordneter" Strukturen, auch Artefakten genannt, wie beispielsweise von atomar definierten Schichtsystemen aus Halbleitern, Isolatoren und/oder Metallen.

In diesem Kapitel werden vor allem die thermischen, elektrischen, magnetischen und optischen Eigenschaften der Festkörper behandelt, die aus der Dynamik der atomaren Schwingungen sowie der Elektronenbewegung im Festkörper zu verstehen sind. Diese Darlegungen schließen sich an die Behandlung des festen Zustands, insbesondere der Kristall- und Realstruktur der Festkörper im vorherigen Kapitel (↑ 6.4.) an; aber auch die Grundlagen des Aufbaus der Atome (↑ 4.2.), der Quantentheorie der chemischen Bindung (↑ 5.1.) und der Statistik von Teilchensystemen (↑ 6.1.) werden für das Verständnis benötigt.

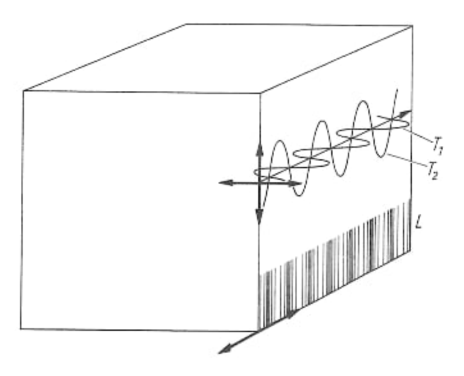


Abb. 8.1.2-1 Ausbreitung von zwei senkrecht zueinander polarisierten transversalen Wellen T_1, T_2 und einer longitudinalen Welle L im Kristall

8.1.2. Gitterschwingungen – Phononen

8.1.2.1. Wesen der Gitterschwingungen. Schwingungen der Atome um ihre Gleichgewichtspositionen bilden die wesentliche Form der thermischen Bewegung im Festkörper. Selbst bei $T = 0$ K befinden sich die Atome nicht völlig in Ruhe, sondern sie führen entsprechend der *Unbestimmtheitsrelation* (s. a. 2.5.1.7.) die sogenannte *Nullpunktsbewegung* aus. Energiezufuhr durch Erwärmen oder auch andere Ursachen (Strahlung) bewirkt ein Anfachen der atomaren Schwingungen, deren Amplituden mit wachsender Temperatur zunehmen. Eine quantitative Behandlung des Problems, wie sie durch BORN und VON KARMAN ab 1910 begründet wurde, muß die Frage beantworten, wieviele verschiedenartige Schwingungsmoden in einem aus N Atomen bestehenden Festkörper möglich sind. Dabei wird jede einzelne Schwingungsmode durch ihre Kreisfrequenz Ω und durch den Wellenzahlvektor $\boldsymbol{q}$ (s. a. 1.2.1.) gekennzeichnet. Die Richtung von $\boldsymbol{q}$ gibt dabei die

Ausbreitungsrichtung der im allgemeinen fortlaufenden Wellen an, die sich aus den Schwingungen der einzelnen Atome infolge ihrer Kopplung durch die Bindungskräfte aufbauen. Der Betrag $|\boldsymbol{q}| = 2\pi/\lambda$ ist der reziproken Wellenlänge proportional. Ferner ist die Angabe der *Polarisation* wichtig, die durch einen Vektor $\boldsymbol{e}$ charakterisiert werden kann. In kubischen, rhombischen oder tetragonalen Kristallen sind, wie Abb. 8.1.2-1 veranschaulicht, longitudinale sowie transversale Schwingungen, letztere in zwei zur Ausbreitungsrichtung senkrechten Polarisationen, möglich. Streng genommen treten rein longitudinale oder transversale Wellen jedoch selbst bei diesen Kristallen nur auf, wenn die Ausbreitungsrichtung mit den Richtungen der Kanten, Flächen- oder Raumdiagonalen der Elementarzellen zusammenfällt.

Einen ersten Einblick in das Wesen der Gitterschwingungen vermittelt das in Abb. 8.1.2-2 dargestellte Modell, bei dem die kugelförmig gedachten Gitterbausteine durch elastische Federn miteinander verbunden sind. Sobald eines dieser „Atome" leicht angestoßen wird, bauen sich infolge der Federkopplung Wellen auf, die sich durch den Modellkristall hindurch fortpflanzen, an den Grenzflächen reflektiert werden und schließlich das gesamte Gebilde in einen Zustand schwingender Bewegung versetzen Die Frage nach der Gesamtzahl voneinander verschiedener Schwingungen läßt sich für dieses Modell leicht beantworten: Nach den Gesetzen der klassischen Mechanik besitzt ein System von N Teilchen insgesamt $3N$ Bewegungsfreiheitsgrade, von denen je drei der Translation beziehungsweise Rotation des gesamten Gebildes entsprechen, während die restlichen $3N - 6$ Freiheitsgrade auf die sogenannten *Normalschwingungen* entfallen.

Identifiziert man nun die Modellkugel mit den atomaren Gitterbausteinen eines Kristalls und betrachtet die Kopplungsfedern als eine Näherung für die atomaren Wechselwirkungskräfte zwischen nächsten Teilchen, so folgt eine wichtige Erkenntnis: In einem Kristall, der aus N gleichartigen atomaren Bausteinen aufgebaut ist, sind insgesamt $3N$ verschiedene Schwingungsmoden möglich. (Wegen der sehr großen Anzahl von Bausteinen in endlichen Kristallen läßt sich der Abzug von 6 Translations- und Rotationsfreiheitsgraden vernachlässigen.) Jede *Schwingungsmode*, auch *Schwingungszustand* genannt, wird durch Angabe der ihr entsprechenden Kreisfrequenz Ω, des Wellenzahlvektors $\boldsymbol{q}$ und der Polarisationsrichtung $\boldsymbol{e}$ gekennzeichnet. Eine wichtige Aufgabe der Gitterdynamik besteht nun darin, diese Daten für das gesamte Spektrum der Gitterschwingungen eines Kristalls zu berechnen. Dabei gelangt man, wie später für eindimensionale Modellfälle noch gezeigt wird (↑ 8.1.2.4.), zu den folgenden Ergebnissen. Die möglichen oder „erlaubten" Schwingungsmoden entsprechen diskreten Eigenwerten des Wellenzahlvektors $\boldsymbol{q}$. Für Wellen längs einer der kristallographischen Achsen-

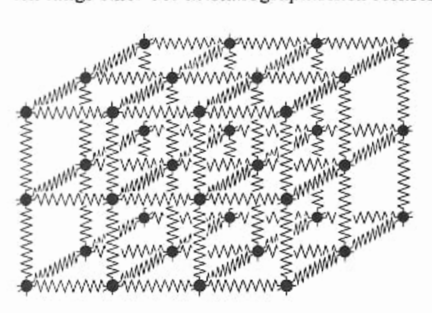


Abb. 8.1.2-2 Federmodell eines schwingenden Festkörpers

richtungen, also im eindimensionalen Fall, findet man die Werte (N gerade)

(2) $q = 0, \pm 2\pi/L, \pm 4\pi/L, \pm 6\pi/L \ldots \pm N\pi/L,$

wenn $L = Na$ (a Gitterkonstante, d. h. Abstand der Gitterpunkte) die Länge des Kristalls in der betrachteten Richtung ist. Die Kreisfrequenz Ω der Wellen, die den durch $\boldsymbol{q}$ gekennzeichneten Schwingungsmoden entsprechen, folgt aus der *Dispersionsrelation*

(3) $\Omega = \Omega(\boldsymbol{q}),$

deren Verlauf im Ergebnis der Rechnungen erhalten wird. Bezeichnen wir die Auslenkung des n-ten atomaren Bausteins aus seiner Gleichgewichtslage mit u_n, so läßt sich die Schwingungsbewegung für die Mode q_j, Ω_j durch die periodische Funktion

(4) $u_n = \hat{u} \exp i(q_j x - \Omega_j t)$ ($\hat{u}$ — Amplitude)

beschreiben, wobei man der Einfachheit halber wiederum auf eine Achsenrichtung, die mit der x-Achse zusammenfallen soll, spezialisiert. Infolge der atomaren Struktur ist diese Funktion nur an den Gitterpunkten $x = na$ definiert. Gleichung 8.1.(4) beschreibt im allgemeinen fortlaufende Wellen, die sich mit der Phasengeschwindigkeit $c_j = \Omega_j/q_j$ ausbreiten. Abb. 8.1.2-3 veranschaulicht die Auslenkung paralleler Gitterebenen beim Durchgang einer longitudinalen bzw. einer transversalen Welle durch den Kristall.

Für Kristalle, bei denen jeder Gitterpunkt nur mit einem Atom besetzt ist, ergibt sich als Dispersionsrelation die Beziehung

(5) $\Omega = \Omega_{\max} |\sin(\frac{1}{2} qa)|,$

deren Verlauf in Abb. 8.1.2-4 dargestellt ist. Im Falle kleiner Wellenzahlen oder großer Wellenlängen läßt sie sich durch die gestrichelt ein-

gezeichnete Gerade $\Omega = c_0 q$ annähern. Für Schall- und Ultraschallwellen mit Frequenzen bis zu 10^8 Hz ist diese Voraussetzung gut erfüllt, und übereinstimmend mit der Erfahrung erhält man die frequenzunabhängige Phasengeschwindigkeit c_0, die allerdings bei dreidimensionalen Kristallen von der Ausbreitungsrichtung und von der Polarisation abhängt. Bei Gitterschwingungen sehr kleiner Wellenlänge treten jedoch die Besonderheiten der atomaren Struktur hervor; die Kreisfrequenz nimmt im Vergleich zum linearen Grenzfall schwächer mit

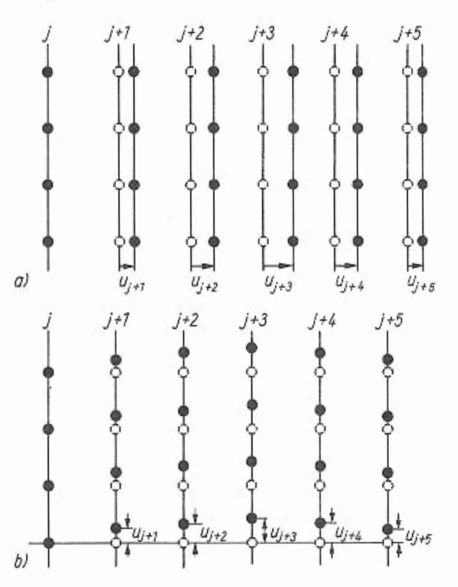


Abb. 8.1.2-3 Auslenkung der Atomebenen im Kristallgitter beim Durchgang a) einer longitudinalen, b) einer transversalen Welle; j Nummer der Atomebene, u_j zugehörige Auslenkung von der Gleichgewichtslage

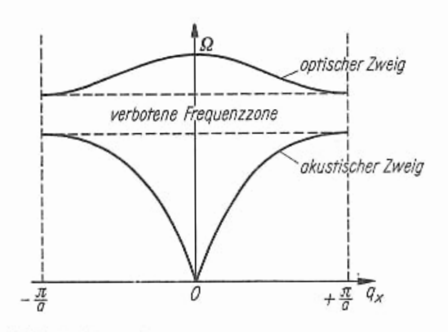


Abb. 8.1.2-4 Schwingungsfrequenzen ω als Funktion der Wellenzahl für einen eindimensionalen Ionenkristall A^+B^- (Erklärung im Text)

der Wellenzahl zu und schließlich wird für $q = \pm\pi/a$ die maximale Kreisfrequenz Ω_{max} erreicht. Sie liegt in der Größenordnung von 10^{13} Hz, und die zugehörige minimale Wellenlänge beträgt $\lambda_{min} = 2a$. Infolge des periodischen Charakters von 8.1.(5) sind noch höhere Wellenzahlen zwar formal möglich, jedoch ergeben sich damit für die durch 8.1.(4) beschriebene atomare Schwingungsbewegung keine neuen Lösungen, sondern die entsprechenden Schwingungsmoden sind in den Lösungen für den Wertebereich $-\pi/a \leqq q \leqq +\pi/a$ bereits enthalten. Der von diesen Grenzen umschlossene Bereich von q wird als *1. Brillouinzone* des – hier zunächst eindimensional betrachteten – Kristalls bezeichnet (↑ 8.1.2.4.). Demnach können in Kristallen nur Gitterschwingungen bis zu einer bestimmten Grenzfrequenz angefacht werden. Wellen dieser Frequenz liegen zwar weit über dem Bereich des Schalls und Ultraschalls, jedoch treten sie bei der Wärmebewegung sowie bei der Wechselwirkung von Lichtquanten oder atomaren Teilchen, wie Elektronen oder Ionen, mit dem Kristallgitter auf. Es sei noch angemerkt, daß der Grenzfall $q = \pm\pi/a$ einer stehenden Welle entspricht, bei der benachbarte Teilchen bzw. Gitterebenen um die Phase π verschoben ausgelenkt werden. Im übrigen sind die hier beschriebenen Gitterschwingungen dadurch gekennzeichnet, daß benachbarte Atome bzw. Gitterebenen gleichsinnig ausgelenkt werden. Da hierzu im langwelligen Bereich die *Schall-* und *Ultraschallwellen* gehören, spricht man vom *akustischen Zweig der Gitterschwingungen.*

Bisher hatten wir nur einen Zweig der akustischen Gitterschwingungen betrachtet. Berücksichtigt man jedoch die drei möglichen Polarisationsrichtungen (eine vorwiegend longitudinale und zwei vorwiegend transversale), so ergeben sich insgesamt drei akustische Zweige mit ähnilchem, jedoch nicht völlig übereinstimmendem Verlauf der Dispersionsrelation. Für einen Kristall mit einem Atom je Gitterpunkt, bzw. einem Atom in der primitiven Elementarzelle (↑ 6.4.1.) ist die Gesamtzahl von $3N$ möglichen Schwingungsmoden damit ausgeschöpft. Enthält die primitive Elementarzelle dagegen mehrere Atome, entfallen also mehrere Atome auf einen Gitterpunkt, treten zusätzlich weitere Schwingungsmoden auf, die man zusammenfassend als *optische Gitterschwingungen* bezeichnet. Bei zwei Atomen je Gitterpunkt, beispielsweise bei Diamant, Silicium oder Natriumchlorid, sind zusätzlich drei Zweige optischer Schwingungsmoden zu berücksichtigen, und allgemein umfaßt das vollständige Schwingungsspektrum eines Kristalls mit p Atomen je Gitterpunkt, bzw. p Atomen in der primitiven Elementarzelle, insgesamt drei Zweige akustische und $3(p - 1)$ Zweige optische Schwingungsmoden. Das Wesen dieser optischen Gitter-

schwingungen soll am Beispiel eines Ionenkristalls der Zusammensetzung A^+B^- erläutert werden. Wie sich rechnerisch zeigen läßt, zerfällt die Dispersionsrelation $\Omega(\boldsymbol{q})$ auch bei Beschränkung auf eine Polarisation in zwei getrennte Zweige, da zu jedem Eigenwert von $\boldsymbol{q}$ zwei verschiedene Kreisfrequenzen erhalten werden. In Abb. 8.1.2-4 sind diese beiden Zweige dargestellt, während Abb. 8.1.2-5 die entsprechenden Schwingungsformen veranschaulicht. Danach schwingen benachbarte Atome bei den Schwingungen des akustischen Zweigs gleich-

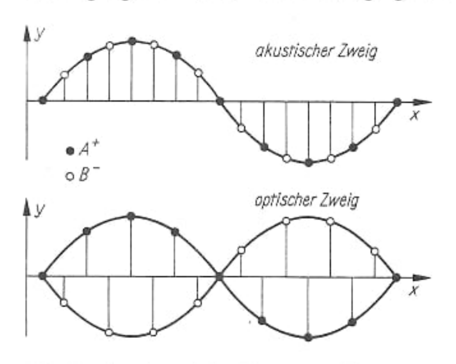


Abb. 8.1.2-5 Typische Bewegungsformen der Atome in einer linearen A^+B^--Kette

sinnig, bei denen des optischen Zweigs dagegen gegenläufig. Eine solche Auslenkung ist zu erwarten, wenn entgegengesetzt geladene Ionen durch die periodisch wechselnde elektrische Feldstärke elektromagnetischer Wellen, beispielsweise bei der Absorption von Licht, ausgelenkt werden, und daraus leitet sich der Name optische Gitterschwingungen ab. Allgemein ist diese Schwingungsform jedoch auch bei neutralen Gitterbausteinen möglich, und optische Schwingungen sind beispielsweise auch an der Wärmebewegung beteiligt. Wir merken an, daß die optischen Zweige von den akustischen Zweigen durch eine Bandlücke oder eine verbotene Zone von Frequenzen getrennt sind, innerhalb der keine stationären Gitterschwingungen auftreten können. Ferner umfassen die optischen Gitterschwingungen nur ein relativ schmales Frequenzband, weshalb sie (besonders in der Umgebung von $\boldsymbol{q} = 0$) häufig näherungsweise als Wellen gleicher Frequenz oder monochromatische Gitterschwingungen angesetzt werden.

Die vorstehend diskutierten Ergebnisse der klassischen Gitterdynamik sind von grundsätzlicher Bedeutung für das Verständnis der mechanischen, elektrischen und optischen Eigenschaften von Festkörpern. Sie zeigen zugleich Besonderheiten der *Wellenausbreitung in Kristallen* auf, die uns bei der Behandlung der Elektronenzustände im Festkörper erneut begegnen werden (↑ 8.2.). Eine quantitative Berechnung der Schwingungsspektren für spezielle Stoffe bereitet vor allem deshalb Schwierigkeiten, weil hierfür die Kräfte zwischen Nachbaratomen im Sinne des einfachen Federmodells nicht ausreichen, sondern die Kenntnis der zwischenatomaren Wechselwirkungspotentiale einschließlich der Kraftwirkungen zwischen weiter voneinander entfernten Gitterbausteinen benötigt wird. Durch Einsatz der elektronischen Rechentechnik konnte dieses Problem zumindest für einfachere Kristalle bewältigt werden. Für das prinzipielle Verständnis wichtiger sind indessen die Konsequenzen, die sich ergeben, wenn man das klassische Modell der Gitterschwingungen den Forderungen der Quantentheorie anpaßt.

8.1.2.2. Quantencharakter der Gitterschwingungen – Phononen. Als gebundene Zustände eines atomaren Systems unterliegen die Gitterschwingungen den Gesetzmäßigkeiten der *Quantentheorie.* Hält man die Nachbaratome im Federmodell zunächst gedanklich fest, so würde sich ein angestoßenes Atom wie ein *harmonischer Oszillator* verhalten. Für die Energiezustände eines solchen Oszillators bei der Schwingung in einer Raumrichtung gilt (↑ 2.5.4.3.)

(6) $E_n = \hbar\Omega(n + 1/2),\ n = 0, 1, 2, \ldots$

Die Energie ist also gequantelt, d. h., die Schwingung des betrachteten Atoms kann nicht kontinuierlich, sondern nur stufenweise verändert werden. Der Abstand zwischen benachbarten Energiezuständen ist durch das Produkt $\hbar\Omega$ gegeben, wobei $\hbar = h/2\pi$ die durch 2π dividierte Plancksche Konstante und Ω die Kreisfrequenz des schwingenden Atoms bezeichnen.

Eine quantenmechanische Behandlung des Problems der kollektiven Gitterschwingungen lehrt nun, daß die wichtige Beziehung 8.1.(6) auch für die im vorigen Abschnitt behandelten Schwingungsmoden gilt. Das Ergebnis der klassischen Berechnung der Schwingungszustände des Festkörpers bleibt also gültig, wenn zusätzlich festgelegt wird, daß jeder dieser Zustände nur Energie im Betrage eines ganzzahligen Vielfachen von $\hbar\Omega$ aufnehmen kann. Eine weitere Besonderheit besteht darin, daß die atomare Bewegung auch für $n = 0$, also beim absoluten Nullpunkt der Temperaturskala und Fehlen jeglicher Energiezufuhr, nicht völlig zur Ruhe kommt. Es verbleibt dann in jeder Schwingungsmode noch die Energie $\hbar\Omega/2$, die auf die von der Quantentheorie geforderte *Nullpunktsbewegung* entfällt.

Der Betrag eines Schwingungs- oder Schallquants läßt sich mit dem Wert der Planckschen Konstante $h = 6{,}62 \cdot 10^{-34}\,\mathrm{Js}^2$ leicht ausrechnen. Für das Quant einer Ultraschallwelle von 10^8 Hz ergibt sich danach eine Energie von

rund $5 \cdot 10^{-7}$ eV. Gemessen an der mittleren thermischen Energie eines Atoms bei Zimmertemperatur (rund $2{,}5 \cdot 10^{-2}$ eV) ist dieser Betrag sehr klein, so daß Quanteneffekte im Bereich der Schall- und Ultraschallwellen meist vernachlässigt werden können. Das gilt aber nicht für die Schwingungen mit viel höheren Kreisfrequenzen (etwa bis zu 10^{13} Hz), die an der thermischen Bewegung oder bei der Wechselwirkung von Elektronen bzw. Lichtquanten mit dem Gitter beteiligt sind.

Die Beziehung $E = \hbar\Omega$ für die Energie eines Schwingungsquants erinnert an die bekannte Formel $E = \hbar\omega$ für die Energie von Lichtquanten oder Photonen (↑ 1.2.2.). Diese Analogie reicht sogar noch weiter. Bei der Wechselwirkung mit Quanten elektromagnetischer Strahlen oder mit atomaren Teilchen verhalten sich die Gitterschwingungen so, als hätten sie einen Impuls von $\hbar\boldsymbol{q}$ ($\boldsymbol{q}$ Wellenzahlvektor der Gitterschwingung). Das ist analog zum Impuls $\hbar\boldsymbol{k}$ eines Lichtquants mit dem Wellenzahlvektor $\boldsymbol{k}$. (Hier wie im folgenden unterscheiden wir die Kreisfrequenzen und Wellenzahlen von Gitterschwingungen zu denen von Lichtquanten durch die unterschiedlichen Bezeichnungen Ω, ω bzw. $\boldsymbol{q}$, $\boldsymbol{k}$.) Die unelastische Streuung von Lichtquanten an Gitterschwingungen, wie sie beim Durchgang von Laserstrahlen durch Kristalle beobachtet wird, läßt sich beispielsweise mit den Gesetzen der *Energieerhaltung* in der Form

$$(7) \quad \hbar\omega = \hbar\omega' \pm \hbar\Omega$$

und der *Impulserhaltung* in der Form

$$(8) \quad \hbar\boldsymbol{k} = \hbar\boldsymbol{k}' \pm \hbar\boldsymbol{q}$$

quantitativ richtig beschreiben. In Abb. 8.1.2-6 sind zwei solche Ereignisse schematisch dargestellt. Der Fall a, für den in den obigen Gleichungen das Pluszeichen gilt, entspricht folgendem Sachverhalt: Das einlaufende Lichtquant

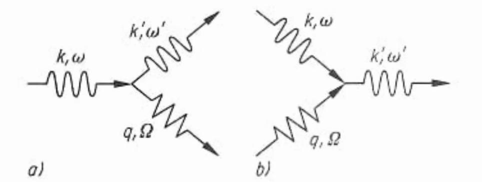


Abb. 8.1.2-6 Photon-Phonon-Streuung; a) Erzeugung, b) Vernichtung eines Phonons $\boldsymbol{q}$

mit der Kreisfrequenz ω und dem Wellenzahlvektor $\boldsymbol{k}$ tritt mit dem Gitter in Wechselwirkung, es wird ein Schwingungsquant mit den Werten Ω, $\boldsymbol{q}$ erzeugt, während das gestreute Quant die kleinere Energie $\hbar\omega'$ und den Wellenzahlvektor $\boldsymbol{k}'$ besitzt. Entsprechend gilt Fall b mit dem Minuszeichen in den Erhaltungssätzen bei Vernichtung eines Schwingungsquants, dessen Energie dem Lichtquant zugeführt wird. Da sich die Schwingungsquanten bei diesen und anderen Prozessen formal wie Teilchen verhalten, ist es üblich, sie als *Phononen* zu bezeichnen.

Ein Phonon oder Schallquant ist also das Energiequant $\hbar\Omega$ einer Gitterschwingung der Kreisfrequenz Ω.

Natürlich stellen die Phononen keine Gebilde dar, die losgelöst vom Festkörper ein Eigendasein führen könnten; sie sind vielmehr als energetische Anregungszustände aufzufassen, die sich unter bestimmten Aspekten wie Teilchen verhalten. Die Phononen bilden ein wichtiges Beispiel für das Konzept der *Quasiteilchen*, das sich auch für die Beschreibung andersartiger energetischer Anregungszustände im Festkörper bewährt hat (↑ 8.2.4.2.). Im übrigen darf die Analogie nicht zu weit geführt werden. So handelt es sich bei dem in Formel 8.1.(8) auftretenden Impuls nicht um einen echten physikalischen Impuls, den die Phononen gar nicht besitzen, sondern um einen *Pseudoimpuls*, der erst zusammen mit der Impulsübertragung an das Gitter des übrigen Festkörpers als Ganzes einen physikalischen Sinn erhält. Genauer ist es daher, Formeln nach der Art von 8.1.(8) nicht als Impulserhaltungssatz, sondern als Auswahlregel für die Wellenzahlvektoren bei Wechselwirkungsprozessen aufzufassen.

Im übrigen hat sich das Phononenkonzept so weitgehend durchgesetzt, daß man die Bezeichnung Phononen gewöhnlich als Synonym für Gitterschwingungen benutzt. Entsprechend wird die oben an Beispielen betrachtete Wechselwirkung zwischen elektromagnetischen Strahlungsquanten und den Gitterschwingungen als *Photon-Phonon-Streuung* bezeichnet. Die analoge Streuung von Neutronen unter Vernichtung oder Erzeugung von Phononen heißt *Neutron-Phonon-Streuung*. Durch genaue Analyse der Frequenz- und Richtungsabhängigkeit des gestreuten Laserlichts bzw. der Energie- und Richtungsverteilung gestreuter Neutronen lassen sich die Schwingungs- oder *Phononenspektren* von Kristallen heute mit großer Genauigkeit experimentell ermitteln.

Weitere wichtige Festkörperprozesse unter Beteiligung von Phononen sind die *Elektron-Phonon-Streuung*, die maßgeblich am Zustandekommen des *elektrischen Widerstands* in Metallen (↑ 8.4.3.) und Halbleitern (↑ 8.3.3.) beteiligt ist, sowie die *Phonon-Phonon-Streuung* (Streuung der Phononen untereinander), die bei der Deutung der *Wärmeleitung* (↑ 8.1.4.) eine wichtige Rolle spielt.

Um die Effekte von Photonen bzw. Phononen in Festkörperprozessen vergleichend einzuschätzen, ist es nützlich, die Energien bei gleicher Wellenzahl oder gleicher Wellenlänge zu betrachten. Mit Werten in der Größenordnung

von $10^3\ \mathrm{ms}^{-1}$ sind die Schallgeschwindigkeiten c_S im Festkörper sehr viel kleiner als die Lichtgeschwindigkeit c_L. Wegen des allgemeinen Zusammenhangs $c_S = \Omega/|\boldsymbol{q}|$ bzw. $c_L = \omega/|\boldsymbol{k}|$ zwischen Phasengeschwindigkeit und Wellenzahl muß daher bei gleicher Wellenzahl $|\boldsymbol{q}| = |\boldsymbol{k}|$ die Kreisfrequenz und somit die Energie des Photons sehr viel größer sein als die des Phonons. Bei den oben betrachteten Prozessen der Photon-Phonon-Streuung wird demzufolge die Energie der Photonen nur wenig verändert, während der Streuwinkel infolge unterschiedlicher Richtung der Vektoren $\boldsymbol{q}$ und $\boldsymbol{k}$ trotz gleichen Betrags der Wellenzahlen beträchtlich sein kann.
Ähnlich gilt für Prozesse unter Beteiligung von Elektronen, daß die Wechselwirkung mit Phononen bei relativ großer Impulsänderung nur kleine energetische Verschiebungen zur Folge hat.

8.1.2.3. Berechnung der Gitterschwingungen für die eindimensionale Atomkette. Um die Auswahl der diskreten Schwingungsmoden mit Hilfe geeigneter Randbedingungen besser zu verstehen und die Aussagen der vorangehenden Abschnitte zu untermauern, wird die Berechnung der Gitterschwingungen nachfolgend am Modellfall eines endlich ausgedehnten *eindimensionalen Kristalls* diskutiert.
Der betrachtete Festkörper soll aus $N + 1$ gleichartigen Atomen der Masse M bestehen, die entlang einer Geraden in gleichen Abständen mit der Gitterkonstanten a angeordnet sind (Abb. 8.1.2-7). Das erste Atom dieser Kette befinde sich in der Ruhelage bei $x = 0$, das letzte bei $x = Na$, wenn x die Ortskoordinate entlang der Kette bezeichnet. Die Variable u soll, wie die Abbildung veranschaulicht, die momentane Verrückung von der Gleichgewichtsposition beschreiben, die entweder longitudinal längs der x-Achse oder senkrecht dazu, also transversal, erfolgen kann. In einem dreidimensionalen kubischen Kristall, bei dem die x-Achse mit einer der Kristallachsen zusammenfällt, würden bei diesen Schwingungen ganze Gitterebenen parallel oder senkrecht zueinander ausgelenkt. Das eindimensionale Modell entspricht dann durchaus den realen Gegebenheiten, wenn man die Atome der Kette als repräsentative Gitterpunkte dieser Ebenen ansieht. Die Endatome bei $x = 0$ und $x = Na$ entsprechen den Oberflächen des Kristalls senkrecht zur x-Achse.
Wir beschränken uns auf die Wechselwirkung zwischen nächsten Nachbaratomen und setzen für die Kraft auf das Atom j der Kette

(9) $F_j = \gamma[(u_{j+1} - u_j) + (u_{j-1} - u_j)]$

Dabei ist γ die Kraftkonstante. Die Kraft soll demnach proportional der Auslenkungen zwischen Nachbaratomen sein, was einer atomistischen Formulierung des Elastizitätsgesetzes von HOOKE entspricht.
Mit diesem Ansatz erhält man für die Bewegungsgleichung des Atoms j

(10) $M\,\mathrm{d}^2u/\mathrm{d}t^2 = \gamma(u_{j+1} + u_{j-1} - 2u_j)$.

Die Summanden auf der rechten Seite bewirken, daß die Bewegungen aller Atome der Kette miteinander gekoppelt sind.
Um 8.1.(10) zu lösen, setzen wir laufende Wellen an, die nur an den Orten $x = ja$ der Atome definiert sein sollen:

(11) $u \;\; = \hat{u}\ \exp\,[\mathrm{i}(jqa - \Omega t)]$,
$\quad\; u_{j+1} = u_j \exp\,[\mathrm{i}(+qa)]$.

Durch Einsetzen in die Differentialgleichung folgt

(12) $\Omega^2 M = -\gamma[\exp(\mathrm{i}qa) + \exp-(\mathrm{i}qa) - 2]$.

Mit Hilfe der mathematischen Identität $\cos x = 1/2[\exp(\mathrm{i}x) - \exp(-\mathrm{i}x)]$ läßt sich dieser Ausdruck umformen in $\Omega^2 = (2\gamma/M)(1 - \cos qa)$.

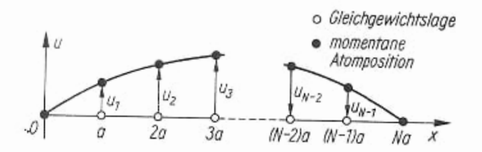


Abb. 8.1.2-7 Auslenkung in einer eindimensionalen Kette von $(N + 1)$ Atomen; die für den transversalen Fall angegebenen Auslenkungen sind übertrieben groß gezeichnet

Durch Anwendung der trigonometrischen Halbwinkelsätze folgt weiter

(13) $\Omega = (4\gamma/M)^{1/2}\,|\sin(\frac{1}{2}\,qa)|$.

Das ist die Dispersionsrelation zwischen Ω und q für die Gitterschwingungen der eindimensionalen Kette. Da der Betrag des Sinus höchstens 1 werden kann, sieht man sofort, daß die Kreisfrequenz für $q = \pm\pi/a$ den Maximalwert $\Omega_{\max} = (4\gamma/M)^{1/2}$ annimmt. Damit haben wir den früher diskutierten Verlauf des Spektrums der akustischen Gitterschwingungen bestätigt (↑ 8.1.2.1., s. a. 8.1.(5) und Abb. 8.1.2-4).
Der Ansatz 8.1.(11) läßt in Verbindung mit der Dispersionsrelation 8.1.(13) zunächst unendlich viele Wellen als Lösung des Problems zu. Um daraus diejenigen auszuwählen, die den endlichen Bewegungsfreiheitsgraden der $N + 1$ Atome der Kette entsprechen, müssen geeignete Randbedingungen auferlegt werden. Zum richtigen Ergebnis führen *periodische Randbedingungen*, nach denen für jede Welle gefordert wird, daß

(14) $u(ja) = u(ja + L)$

erfüllt ist. Der linke und der rechte Rand der Kette (oder in drei Dimensionen des Kristalls) schwingen dann phasengleich, und die gesamte Gitterschwingung wiederholt sich in Intervallen der Länge L. Physikalisch ist die Wahl dieser Randbedingungen damit zu rechtfertigen, daß die Lösungen unverändert bleiben, wenn man zu Ketten bzw. Kristallen mehrfacher Länge übergeht.

Man überzeugt sich leicht, daß die Wellen des Lösungsansatzes 8.1.(11) den Randbedingungen genügen, falls die Wellenzahl einen der folgenden Werte annimmt (N gerade):

$$(15)\quad q = 0,\ \pm 2\pi/L,\ \pm 4\pi/L,\ \pm 6\pi/L, \dots \pm N\pi/L.$$

Hier entsprechen negative Wellenzahlen solchen Lösungen, die sich bezüglich der x-Achse nach rechts fortpflanzen, während das positive Zeichen für entgegengesetzt laufende Wellen gilt. Speziell für die maximale Wellenzahl $|q| = N\pi/L = N\pi/Na = \pi/a$ läßt sich zeigen, daß die Lösung die Schwingungsform einer stehenden Welle annimmt.

Die Ergebnisse der linearen Atomkette können leicht auf die Verhältnisse eines dreidimensionalen kubischen Kristalls erweitert werden. Der würfelförmige Kristall mit der Kantenlänge L soll mit seinen Achsen parallel zu den Koordinatenachsen eines rechtwinkligen Systems x, y, z angeordnet sein. Eine ebene Welle in dem Kristall wird beschrieben durch

$$(16)\quad u = \hat{u} \exp i(\boldsymbol{qr} - \Omega t) = \hat{u}[\exp i(q_x x + q_y y + q_z z) \cdot \exp(-i\Omega t)]$$

mit den Komponenten q_x, q_y, q_z des Wellenzahlvektor . Von diesen Wellen sind als erlaubte Schwingungszustände des Kristalls diejenigen auszuwählen, die den dreifach periodischen Randbedingungen

$$(17)\quad \exp[i(q_x x + q_y y + q_z z) = \exp[i(q_x(x + L) + q_y(y + L) + q_z(z + L)]$$

genügen. Diese Bedingungen sind erfüllt, wenn die Komponenten von $\boldsymbol{q}$ analog zu 8.1.(15) die folgenden Werte annehmen (N gerade):

$$(18)\quad \left.\begin{matrix} q_x \\ q_y \\ q_z \end{matrix}\right\} = 0,\ \pm 2\pi/L,\ \pm 4\pi/L,\ \pm 6\pi/L, \dots \pm N\pi/L.$$

N ist die Anzahl der Atome in Richtung einer Kristallachse, so daß die Gesamtzahl der Atome des Kristalls $\hat{N} = N^3$ beträgt. Jedem Zahlentripel der Komponenten von $\boldsymbol{q}$ entsprechen drei Schwingungsmoden, und zwar eine longitudinale sowie zwei senkrecht zueinander polarisierte transversale Wellen.

Durch Energiezufuhr, beispielsweise im Verlauf der Wärmebewegung oder als Folge der Einwirkung von Strahlen, kann jede dieser Schwingungsmoden in der Weise angeregt werden, daß die Bewegungsenergie gegenüber der Nullpunktsbewegung um ganzzahlige Vielfache von $\hbar\Omega$ erhöht wird. In der Sprache des Quasiteilchenmodells können wir dies auch so ausdrücken: Jede Schwingungsmode kann mit einer ganzen Zahl von Phononen besetzt sein.

Erinnert sei nochmals daran, daß diese *akustischen Phononen* nur für Kristalle mit einem Atom pro Gitterpunkt das vollständige Schwingungsspektrum beschreiben. Bei Kristallen mit mehreren Atomen oder Ionen pro Gitterpunkt sind zusätzlich die *optischen Phononen* zu berücksichtigen, deren Energie wegen der höheren Kreisfrequenzen im optischen Zweig der Schwingungsspektren größer ist als die der akustischen Phononen.

8.1.2.4. Reziprokes Gitter – Brillouinzonen. Die relativ komplizierten Erscheinungen der Wellenausbreitung in Kristallen lassen sich in übersichtlicher Weise darstellen, indem man die Hilfsvorstellung des *reziproken Gitters* einführt.

Die Punkte des gewöhnlichen dreidimensionalen Raumgitters sind festgelegt durch (↑ 4.1.1.);

$$(19)\quad \boldsymbol{R} = n_1\boldsymbol{a} + n_2\boldsymbol{b} + n_3\boldsymbol{c},$$

wobei $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ die *fundamentalen Translationsvektoren* oder Gittergrundvektoren und n_1, n_2, n_3 ganze Zahlen sind. $\boldsymbol{R}$ ist dann ein *Gittervektor*, der bei gegebenen Maßzahlen zwei Gitterpunkte verbindet. Die *fundamentalen Translationsvektoren des zugehörigen reziproken Gitters* werden nun durch die folgenden Gleichungen definiert:

$$(20)\quad \boldsymbol{a} = 2\pi(\boldsymbol{b} \times \boldsymbol{c})/[\boldsymbol{abc}];\ \boldsymbol{b} = 2\pi(\boldsymbol{c} \times \boldsymbol{a})/[\boldsymbol{abc}];\ \boldsymbol{c} = 2\pi(\boldsymbol{a} \times \boldsymbol{b})/[\boldsymbol{abc}].$$

Hier steht das Zeichen $\times$ für das Vektorprodukt, und mit $[\boldsymbol{abc}] = \boldsymbol{a}(\boldsymbol{b} \times \boldsymbol{c})$ wird in üblicher Weise das Spatprodukt bezeichnet. Da das Spatprodukt verschwindet, wenn zwei der Vektoren gleich sind, folgt für die skalaren Produkte der Translationsvektoren des gewöhnlichen und des reziproken Gitters:

$$\begin{aligned} &\boldsymbol{a}^*\boldsymbol{a} = 2\pi; \quad &\boldsymbol{ab}^* = 0\ ; \quad &\boldsymbol{ac}^* = 0\ ; \\ &\boldsymbol{ba}^* = 0\ ; \quad &\boldsymbol{bb}^* = 2\pi; \quad &\boldsymbol{bc}^* = 0\ ; \\ &\boldsymbol{ca}^* = 0\ ; \quad &\boldsymbol{cb}^* = 0\ ; \quad &\boldsymbol{cc}^* = 2\pi. \end{aligned}$$

Geometrisch bedeuten diese Beziehungen, daß ein Translationsvektor des reziproken Gitters auf jeweils zwei Translationsvektoren des gewöhnlichen Gitters senkrecht steht, und natürlich gilt auch das Umgekehrte.

Aus den fundamentalen Translationsvektoren $\boldsymbol{a}^*, \boldsymbol{b}^*, \boldsymbol{c}^*$ des reziproken Gitters läßt sich eine Elementarzelle bilden, deren Gestalt im allgemeinen von der des gewöhnlichen Gitters verschieden ist. Nur im Falle des einfach kubischen Gitters stimmen die Elementarzellen des gewöhnlichen und des reziproken Gitters gestalt-

mäßig überein, jedoch ist die Gitterkonstante des reziproken Gitters umgekehrt proportional zu der des gewöhnlichen Gitters. Als weitere Beispiele führen wir an, daß die reziproke Elementarzelle des kubisch flächenzentrierten Gitters kubisch raumzentriert ist, während sich umgekehrt für das kubisch raumzentrierte Gitter in der reziproken Darstellung eine kubisch flächenzentrierte Elementarzelle ergibt.
Weiter können wir analog zu 8.1.(19) den allgemeinen *Gittervektor G des reziproken Gitters* angeben: ($\hat{h}, \hat{k}, l$ ganze Zahlen).

(21) $G = \hat{h}a^* + \hat{k}b^* + lc^*$.

Der so definierte Vektor G verbindet für gegebene Maßzahlen zwei Punkte des reziproken Gitters. Legt man den Ausgangspunkt in den Ursprung des Koordinatensystems und läßt die Maßzahlen $\hat{h}$, $\hat{k}$, l der Reihe nach alle ganzen Werte durchlaufen, so beschreibt 8.1.(21) im Sinne der allgemeinen Gittertranslation die Lage aller Punkte des reziproken Gitters.
Die Gitterebenen, die den Nullpunkt des reziproken Gitters umgeben, lassen sich zu geschlossenen räumlichen Zellen zusammenfügen, die man *Brillouinzonen* nennt. Am Beispiel der eindimensionalen Atomkette sowie eines zweidimensionalen quadratischen Gitters soll dies näher betrachtet werden (↑ Abb. 8.4.4-3). Für das eindimensionale Gitter mit dem Punktabstand a sind die kleinsten reziproken „Vektoren" durch die Strecken $\pm 2\pi/a$ gegeben, der allgemeine „Vektor" des reziproken Gitters ist ein ganzes Vielfaches dieser fundamentalen Translation. Unterteilt man diese Strecken jeweils im halben Abstand vom Nullpunkt, so ergeben sich Bereiche, die der Reihe nach als 1., 2. usw. Brillouinzone bezeichnet werden. Die 1. Brillouinzone erstreckt sich von $-\pi/a$ bis $+\pi/a$, während die 2. Brillouinzone aus den beiden Abschnitten zwischen $-2\pi/a$ und $-\pi/a$ sowie zwischen $+\pi/a$ und $+2\pi/a$ besteht. Die Nützlichkeit der Einführung des reziproken Gitters zeigt sich nun, wenn man die Wellenzahlen q der erlaubten Schwingungsmoden gemäß 8.1.(15) einträgt, also die Achse des eindimensionalen reziproken Gitters als Koordinate für die Wellenzahlen $q = 2\pi/\lambda$ nutzt, die ja gleichfalls die Dimension einer reziproken Länge besitzen. Die höchste Wellenzahl der akustischen Phononen fällt dann mit den Grenzen der 1. Brillouinzone zusammen, während die übrigen Moden innerhalb dieser Zone gleichabständig in Intervallen $q = 2\pi/L$ angeordnet sind. Die *Zustandsdichte* Z, d. h. die Anzahl der Moden pro Längeneinheit der q-Achse beträgt demnach $L/2\pi$ innerhalb der 1. Brillouinzone und 0 für Wellenzahlen außerhalb davon.
Um die Einteilung in Brillouinzonen bei einem zweidimensionalen quadratischen Gitter der Gitterkonstanten $2\pi/a$ vorzunehmen, bildet man auf den kürzesten Gittervektoren die Mittelsenkrechten und verlängert sie bis zu den Schnittpunkten. Als 1. bis 3. Brillouinzonen werden dann die Flächenabschnitte erhalten, die in Abb. 8.1.2-8 verschieden schraffiert dargestellt sind. Die 1. Zone ist ein Quadrat, aus dem beim Übergang zum dreidimensionalen, einfach kubischen Gitter ein Würfel hervorgeht.
Trägt man wiederum die Wellenzahlen der erlaubten Schwingungsmoden für alle drei Komponenten entsprechend 8.1.(18) ein, so ergibt sich, daß die Schwingungsmoden gerade die 1. Brillouinzone ausfüllen. Da der Abstand der Moden längs jeder Achse $2\pi/L$ beträgt, entfällt eine Schwingungsmode auf das Volumen $(2\pi/L)^3$ im Raum des reziproken Gitters. Dieser Raum der reziproken Längen wird als *Fourierraum* oder auch als $\boldsymbol{q}$-Raum bzw. $\boldsymbol{k}$-Raum (nach den Symbolen für die Wellenzahlvektoren von Phononen bzw. Photonen) bezeichnet. Die Zustandsdichte, d. h. die Zahl der Schwingungsmoden pro Volumeneinheit des Fourierraums, beträgt

(22) $Z(\boldsymbol{q}) = (L/2\pi)^3 = V/8\pi^3$,

für das Innere der 1. Brillouinzone und ist außerhalb davon Null ($V = L^3$ Kristallvolumen).
Die hier an einfachsten Gittern diskutierten Ergebnisse lassen sich verallgemeinern. Bei komplizierteren Gittern nimmt die 1. Brillouinzone die Gestalt eines Polyeders an. Innerhalb dieses Bereichs verteilen sich die Schwingungsmoden gleichmäßig im $\boldsymbol{q}$-Raum, so daß die Ermittlung der für viele Erscheinungen wichtigen Zustandsdichte auf eine Volumenberechnung hinausläuft. Betont sei nochmals, daß Wellenzahlvektoren, die außerhalb der 1. Brillouinzone liegen, keinen neuen physikalischen Sachverhalt verkörpern. Durch Addition oder Subtraktion eines passenden reziproken Gittervektors $\boldsymbol{G}$ läßt sich stets eine *Reduktion* in der Weise vornehmen, daß der resultierende Wellenzahlvektor $\boldsymbol{q}' = \boldsymbol{q} \pm \boldsymbol{G}$ in die 1. Brillouinzone fällt. Wegen der Periodizität der Gittereigenschaften ist die zugehörige Wellenbewegung der Atome, wie allgemein bewiesen werden kann, in den Lösungen für die erlaubten $\boldsymbol{q}$-Werte der 1. Brillouinzone bereits enthalten.

8.1.2.5. Photonenstreuung im Kristall. Mit den im vorangehenden Abschnitt erarbeiteten Vorstellungen sollen die Prozesse der Streuung von Photonen, speziell von Quanten der Röntgenstrahlen, in Kristallen etwas eingehender analysiert werden. Betrachtet sei zunächst die elastische Streuung von Photonen, bei der die Energie $\hbar\omega$ sowie der Betrag $|\boldsymbol{k}|$ des Wellenzahlvektors sich nicht ändern. Die Erfahrung zeigt nun, daß beim Einfall eines fein ausgeblendeten Röntgenstrahls einheitlicher Quantenenergie in einen idealen Kristall abgebeugte Strahlen gestreuter Quanten nur in solchen Richtungen

beobachtet werden, für die folgende Bedingung erfüllt ist:

(23) $\boldsymbol{k}' = \boldsymbol{k} + \boldsymbol{G}$.

Hier bezeichnen $\boldsymbol{k}$ bzw. $\boldsymbol{k}'$ die Wellenzahlvektoren des einlaufenden bzw. des abgebeugten Strahls, während $\boldsymbol{G}$ ein beliebiger Vektor sein kann, der zwei Gitterpunkte des reziproken Gitters verbindet. Bei der elastischen Streuung gilt außerdem $|\boldsymbol{k}'| = |\boldsymbol{k}|$, d. h., die Wellenzahlvektoren müssen die gleiche Länge besitzen. Dies führt auf die *Ewaldsche Konstruktion* zum Auffinden der Richtungen, in denen ein abgebeugter Strahl auftreten kann (Abb. 8.1.2-8). Man trägt den Wellenzahlvektor $\boldsymbol{k}$ im Raum des reziproken Gitters so ab, daß er an einem reziproken Gitterpunkt P endet. Bildet man nun eine Kugel mit dem Radius $|\boldsymbol{k}| = \overline{OP}$ um den Nullpunkt, so müssen die möglichen $\boldsymbol{k}'$ gleichfalls Radienvektoren dieser *Ewaldkugel* sein. Davon liefern wegen Gleichung 8.1.(23) jedoch nur diejenigen $\boldsymbol{k}'$ einen abgebeugten Strahl, die gleichfalls an einem Gitterpunkt enden. Elastisch gestreute Strahlen, sogenannte *Beugungsreflexe*, werden daher nur beobachtet, wenn Gitterpunkte des reziproken Gitters auf der Ewaldkugel liegen. Führen wir den Winkel 2Θ zwischen $\boldsymbol{k}$ und $\boldsymbol{k}'$ ein, so folgt aus Abb. 8.1.2-8 als Bedingung zwischen den Beträgen von $\boldsymbol{G}$ und $\boldsymbol{k}$

(24) $|\boldsymbol{G}| = 2\,|\boldsymbol{k}| \sin \Theta$.

Mit Hilfe weiterführender geometrischer Überlegungen, die wir hier überspringen, läßt sich zeigen, daß diese Formel mit der *Braggschen Gleichung* $n\lambda = 2d \sin \Theta$ identisch ist, die bei der Behandlung der Strukturuntersuchung von Kristallen mittels Röntgenstrahlen eingeführt worden ist (↑ 6.4.2.2.1.). Die vorstehend durchgeführte Interpretation liefert eine wichtige,

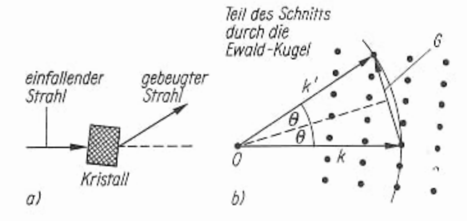


Abb. 8.1.2-8 Röntgenbeugung im Kristall; a) Schema des Strahlengangs, b) Ewaldsche Konstruktion

weiterführende Erkenntnis: Die Beugungsbilder, die man bei der Strukturuntersuchung von Festkörpern mit Röntgenstrahlen erhält, stellen letztlich eine *Abbildung der Punkte des reziproken Gitters* dar. Das trifft in gleicher Weise auch für Beugungsdiagramme zu, die mit Elektronen- oder Neutronenstrahlen erhalten werden, wie man sofort sieht, wenn $\boldsymbol{k}$ und $\boldsymbol{k}'$ als Wellenzahlvektoren von Materiewellen aufgefaßt werden. Multipliziert man Gleichung 8.1.(23) auf beiden Seiten mit $\hbar$, so nimmt sie die Form des Impulssatzes an, da ja $\hbar\boldsymbol{k}$ der Impuls eines Photons mit dem Wellenzahlvektor $\boldsymbol{k}$ ist. Das Glied $\hbar\boldsymbol{G}$ ist dann als derjenige Impulsanteil aufzufassen, der bei der elastischen Streuung an das Gitter übertragen wird.

Nun befinden sich die Kristallproben bei der Strukturuntersuchung mit Röntgenstrahlen bzw. Elektronen- oder Neutronenstrahlen nicht am absoluten Nullpunkt, sondern gewöhnlich werden die Messungen bei Zimmertemperatur durchgeführt. Neben der elastischen Streuung der Photonen ist dann eine *unelastische Wechselwirkung* mit den Phononen zu berücksichtigen. Um diesen Effekt zu erfassen, ist Gleichung 8.1.(23) um ein Glied $\boldsymbol{q}$ zu erweitern, welches dem *Pseudoimpuls* eines Phonons entspricht [↑ 8.1.2.2. sowie 8.1.(8)]:

(25) $\boldsymbol{k}' = \boldsymbol{k} \pm \boldsymbol{q} + \boldsymbol{G}$.

Hier gilt das Minuszeichen, wenn bei dem Prozeß ein Phonon mit dem Wellenzahlvektor $\boldsymbol{q}$ erzeugt wird. Die entsprechende Gitterschwingung wird dann auf Kosten der Energie und des Impulses des Röntgenquants angefacht. Umgekehrt kann auch ein Phonon absorbiert werden, was eine Zunahme der Energie und des Wellenzahlvektors der Röntgenquanten bewirkt, und in diesem Fall ist in obiger Gleichung das Pluszeichen anzusetzen. Diese Formen der Wechselwirkung mit den Gitterschwingungen werden bei der *Strukturuntersuchung mit Röntgen-, Elektronen- oder Neutronenstrahlen* tatsächlich beobachtet. Wegen der dichten Folge möglicher Phononenzustände äußern sie sich durch einen diffusen Anteil in den Beugungsdiagrammen, dessen Intensität mit steigender Temperatur der Festkörperprobe zunimmt. Besonders bei der Beugung von Neutronenstrahlen in Kristallen ist es möglich, die unelastische Streuung durch genaue Analyse der Energie bzw. des Impulses der gestreuten Teilchen quantitativ zu erfassen. Auf diese Weise wird die Gestalt der Phononenspektren von Festkörpern experimentell ermittelt.

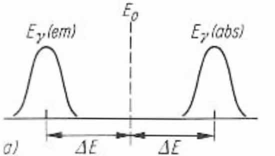
Im übrigen kommt es infolge der Wärmebewegung nicht etwa, wie man zunächst erwarten könnte, zu einer Verbreiterung der Beugungsreflexe, sondern die Braggmaxima der elastischen Streuung bleiben durchaus scharf, jedoch nimmt ihre Intensität mit wachsender Temperatur ab. Nach Debye läßt sich dies folgendermaßen erklären: Zu den Reflexen der elastischen Streuung können nur die Prozesse beitragen, bei denen, gewissermaßen in Momentaufnahme, die strenge Gitterperiodizität erfüllt ist, so daß sich die Phasen der abgebeugten Teil-

wellen konstruktiv überlagern. Daneben werden mit wachsender Temperatur in zunehmendem Maße infolge der Photon-Phonon-Wechselwirkung verstimmte Quanten gestreut, die in der Gesamtheit den diffusen Untergrund der Beugungsdiagramme erzeugen. Quantitativ läßt sich die Intensitätsabnahme der Braggreflexe durch den *Debye-Waller-Faktor* beschreiben, der exponentiell von der Temperatur sowie vom Einfallswinkel und der Wellenlänge der Röntgenstrahlen abhängt. Die Photon-Phonon-Wechselwirkung von Licht im sichtbaren Spektralgebiet bzw. im Infraroten sowie die Elektron-Phonon-Wechselwirkung von Leitungselektronen sind äußerst bedeutungsvoll für das Verständnis *optisch induzierter Elektronenübergänge* (↑ 8.8.3.) sowie für die Deutung der *elektrischen Leitfähigkeit* in Festkörpern (↑ 8.2.3.6.). In Erweiterung von Gleichung 8.1.(24) sind dabei auch *Multiphononenprozesse* zu beachten, bei denen gleichzeitig mehrere Phononen beteiligt sind.

8.1.2.6. Mößbauereffekt. Dieser Effekt, der 1958 von MÖSSBAUER nachgewiesen wurde, stellt eine interessante Verknüpfung zwischen Atomkernprozessen und Vorgängen im Festkörper dar; er verdeutlicht die Kopplung von Erscheinungen in atomaren Dimensionen infolge der Gesetzmäßigkeiten der Quantentheorie.

Wenn ein Atomkern beim radioaktiven Zerfall oder im Gefolge einer Kernreaktion (s. a. 3.4.1. und 3.4.2.) ein γ-Quant der Energie $E_\gamma = \hbar\omega$ mit dem Impuls $p_\gamma = \hbar\omega/c = E_\gamma/c$ (c Lichtgeschwindigkeit) emittiert, so muß der Impulssatz erfüllt sein. Im Falle einzelner, freier Atome erfährt das emittierende Atom daher einen Rückstoß mit dem Impuls E_γ/c. Die entsprechende kinetische Energie beträgt $E_\gamma^2/2Mc^2$ (M Masse des Atoms), da ja zwischen kinetischer Energie und Impuls die allgemeine Beziehung $E_k = p^2/2m$ besteht (s. a. 2.1.3.). Diese Rückstoßenergie muß natürlich auf Kosten der Umwandlungsenergie E_0 gehen, die dem Übergang zwischen zwei scharfen Anregungszuständen des Atomkerns entspricht. Für das ausgesendete γ-Quant verbleibt daher nur die etwas kleinere Energie $E_\gamma = E_0 - E_\gamma^2/2Mc^2$. Bei dem Nuklid Eisen-57 ($^{57}_{26}$Fe) beträgt die Differenz $\Delta E = E_0 - E_\gamma \approx 5 \cdot 10^{-3}$ eV; bezogen auf die Umwandlungsenergie $E_0 = 14{,}4$ keV ergibt sich als relative Verstimmung $\Delta E/E_0 \approx 3 \cdot 10^{-7}$. Da die Eigenbreite der Kernzustände größenordnungsmäßig nur 10^{-13} ist, können die verstimmten Quanten beim Auftreffen auf Eisenkerne im Grundzustand nicht absorbiert werden, denn die für eine *Resonanzabsorption* notwendige Übereinstimmung der Quantenenergie und der Energiedifferenz zum angeregten Zustand ist nicht erfüllt (Abb. 8.1.2-9).

Der Mößbauereffekt besteht nun darin, daß die Verstimmung der γ-Quanten weitgehend vermieden werden kann, wenn die emittierenden Atome starr in das Gitter eines Festkörpers eingebaut sind. Dies ist verständlich, weil der Rückstoß dann nicht mehr auf ein einzelnes Atom, sondern auf den Kristall als Ganzes übertragen wird. In die obigen Ausdrücke ist für diesen Fall als Masse M die Gesamtmasse des Kristalls einzusetzen, die selbst für kleine Proben relativ zur Atommasse schon so groß ist,



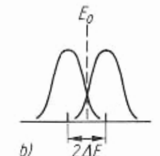


Abb. 8.1.2-9 Mößbauereffekt; a) Rückstoßverstimmung zwischen emittierten und absorbierten γ-Quanten, b) Resonanzfluoreszenz bei Einbau der Emitter- und Absorberatomkerne im Festkörper

daß die Rückstoßenergie auf unmeßbar kleine Werte absinkt. Praktisch wird jedoch beobachtet, daß mit erheblicher Wahrscheinlichkeit ein Energieaustausch mit den Phononen, beispielsweise ein Anfachen bestimmter Gitterschwingungen, erfolgt. Um einen wesentlichen Bruchteil wirklich unverstimmter Quanten zu erhalten, muß man die Festkörperproben auf niedrige Temperaturen kühlen. Das geschieht mit flüssigem Stickstoff oder sogar mit Wasserstoff bzw. Helium. Günstige Bedingungen liegen vor, wenn die Energie der γ-Quanten relativ klein ist und der Festkörper infolge fester Bindung eine möglichst hohe mittlere Phononenfrequenz oder *Debyetemperatur* (↑ 8.1.3.4.) aufweist. Dies ist bei dem schon erwähnten Nuklid Fe-57 der Fall. Man erzeugt geeignete Festkörperproben durch Bestrahlen von metallischem Eisen mit Deuteronen. Dabei bildet sich nach einer (d,n)-Reaktion das Nuklid Co-57, bei dessen Zerfall die gewünschten 14,4 keV-Quanten von Fe-57 ausgestrahlt werden. Von derartigen Proben erhält man selbst bei Zimmertemperatur noch 70% Quanten ohne Rückstoßverstimmung. Bringt man in den Strahlenweg eine inaktive Eisenprobe, so werden diese Quanten mit erheblicher Wahrscheinlichkeit absorbiert, indem Fe-57-Kerne in den angeregten Zustand überführt werden. Allerdings beträgt der Gehalt von gewöhnlichem Eisen an stabilem Fe-57 nur 2,17%, weshalb für empfindliche Messungen Proben mit Isotopenanreicherung eingesetzt werden. Außerdem ist zu beachten, daß auch bei der Absorption durch Impulsübertrag in Vorwärtsrichtung eine Verstimmung im gleichen

Betrag wie bei der Emission erfolgen kann. Hinsichtlich Beschaffenheit der Proben, Kühlung usw. gelten daher für den Absorber die gleichen Anforderungen wie für den Emitter. Die Resonanz-Absorptions-Spektroskopie von γ-Strahlen, meist *Mößbauerspektroskopie* genannt, läßt sich heute mit Nukliden vieler Elemente durchführen. Wegen der außerordentlichen Resonanzschärfe der Atomkernzustände eignet sich diese Methode für hochempfindliche Präzisionsmessungen vieler physikalischer Größen, z. B. magnetischer oder elektrischer Felder am Ort der Kerne. Selbst die geringe Verstimmung von γ-Quanten durch das Schwerefeld der Erde, wie sie von EINSTEIN als Rotverschiebung elektromagnetischer Strahlen in Gravitationsfeldern vorausgesagt wurde, konnte mit einer Genauigkeit von rund 1% bestätigt werden.

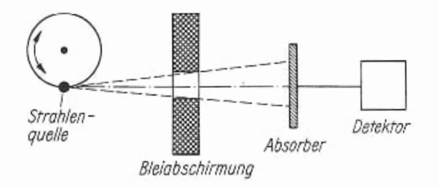


Abb. 8.1.2-10 Mößbauereffekt bei bewegter Strahlenquelle

Verblüffende Ergebnisse erhielt bereits MÖSSBAUER, als er Quelle und Absorberprobe relativ zueinander bewegte, was durch Rotation oder Schwingung geschehen kann (Abb. 8.1.2-10). Dabei erfolgt eine geringfügige Verstimmung durch den *Dopplereffekt* (↑ 2.3.1.9.), die bezogen auf die Quantenenergie $\Delta E = vE/c$ (v Relativgeschwindigkeit) ausmacht. Bereits bei einer relativen Bewegung von wenigen Millimetern pro Minute ist häufig schon ein deutlicher Einfluß auf die Absorption der Quanten festzustellen. Im übrigen treten Dopplerverschiebungen bei Festkörperproben oberhalb des absoluten Nullpunkts durch die Wärmebewegung auf, was sich als Untergrund in den bei variierender Relativgeschwindigkeit erhaltenen Spektren äußert.

8.1.3. Wärmekapazität der Festkörper

8.1.3.1. Empirisches Verhalten. Die innere Energie U eines Festkörpers setzt sich aus der potentiellen und der kinetischen Energie der atomaren Bausteine zusammen, die im wesentlichen nur Schwingungen um ihre Gleichgewichtslagen ausführen können. Als *Wärmekapazität* C_v wird der partielle Differentialquotient der inneren Energie U nach der Temperatur T bei konstantem Volumen V definiert, d. h. $C_v \equiv (\partial U/\partial T)_v$. Zum Vergleich der kalorischen Eigenschaften verschiedener Stoffe bezieht man die Wärmekapazität auf die Masseeinheit. Je nachdem, ob dabei die gewöhnliche Masse m oder die Molmasse M benutzt wird, bezeichnet man die so gewonnenen Stoffkenngrößen als *spezifische Wärmekapazität* c_v oder als *molare Wärmekapazität* c_{vm} (s. a. 6.1.2.). Die entsprechenden Maßeinheiten sind $\mathrm{Jkg^{-1}K^{-1}}$ bzw. $\mathrm{Jmol^{-1}K^{-1}}$; daneben findet man noch häufig die älteren Angaben, bei denen die Energie statt in Joule (J) in Kalorien (cal) gemessen wird.

Experimentell wird bei festen Stoffen allerdings fast immer die spezifische Wärmekapazität bei konstantem Druck c_p bzw. c_{pm} bestimmt, weil sich ein konstantes Volumen nur mit Hilfe sehr hoher Drücke aufrechterhalten ließe. Mit Hilfe der Beziehung $c_{pm} - c_{vm} = V_m T\alpha^2/\varkappa$ läßt sich c_{vm} bei Kenntnis des Molvolumens V_m, des kubischen Ausdehnungskoeffizienten α und der Kompressibilität $\varkappa$ aus den Meßwerten für c_p berechnen (s. a. 8.1.4.4.). Zur Messung von c_p dienen Kalorimeter, in denen die Temperaturänderung der Probe nach Zufuhr einer genau bekannten Wärmemenge bestimmt wird. Moderne Geräte erlauben die Messung von c_p als Funktion der Temperatur in weiten Bereichen, besonders auch bei tiefen Temperaturen.

Zahlenwerte für die bei 25 °C gemessenen molaren Wärmekapazitäten einer Reihe fester Elemente und Verbindungen sind in Tabelle 8.1.3-1 zusammengestellt. Man entnimmt, daß sich

Tab. 8.1.3-1 Molare Wärmekapazitäten c_p und c_v (in $\mathrm{Jmol^{-1}K^{-1}}$) einiger Elemente und Verbindungen bei 25 °C

Stoff	c_p	c_v
Li	23,64	22,53
Na	28,12	24,58
K	29,51	24,80
Mg	24,81	23,46
Ca	26,28	24,21
Al	24,34	23,03
Cu	24,47	23,63
Ag	25,49	24,27
Mo	23,75	23,06
Pb	26,82	24,84
W	24,08	23,59
NaCl	50,79	48,50
KBr	52,49	51,51
$CaCO_3$	83,50	82,52
Al_2O_3	79,01	78,96
SiO_2	44,48	41,06
AgCl	55,78	52,48
$FeCO_3$	82,10	78,81

c_{vm} und c_{pm} jeweils nur um wenige Prozent unterscheiden. Weiter fällt auf, daß die molaren Wärmekapazitäten der aufgeführten Elemente c_{vm} ähnliche Werte aufweisen, die etwas unterhalb von 25 $\mathrm{J mol^{-1} K^{-1}}$ liegen. Diese auffallende Übereinstimmung wurde schon im Jahre 1819 bemerkt und nach den Entdeckern *Dulong-Petitsche Regel* genannt. Wie viel später erkannt wurde, folgt sie aus dem *Gleichverteilungssatz der statistischen Mechanik* (↑ 6.1.3.), nach dem auf jeden Schwingungsfreiheitsgrad pro Atom der mittlere Energiebetrag $2k_BT/2$ (k_B Boltzmannkonstante) oder pro Mol RT (R molare Gaskonstante) entfallen muß. Multipliziert man wegen der Schwingungsmöglichkeiten in den verschiedenen Raumrichtungen mit dem Faktor 3, so ergibt sich nach Division durch T als Wert für die molare Wärmekapazität ungefähr der Betrag von 25 $\mathrm{J mol^{-1} K^{-1}}$. Bei einigen fest gebundenen Kristallen mit niederer Atommasse, insbesondere Diamant, werden allerdings bei Zimmertemperatur viel kleinere molare Wärmekapazitäten gemessen ,die sich erst bei sehr hohen Temperaturen dem von der Dulong-Petitschen Regel geforderten Wert annähern. Auch die *Neumann-Koppsche Regel*, nach der sich die molaren Wärmekapazitäten von Verbindungen additiv aus den Beiträgen der darin enthaltenen Elemente zusammensetzen, wird vielfach nur mangelhaft befolgt. Schließlich weist der Temperaturverlauf der spezifischen Wärmekapazitäten mit sinkender Temperatur einen allmählichen Abfall auf Null auf (Abb. 8.1.3-2). Im Bereich tiefer Temperaturen steigt c_v etwa proportional zur dritten Potenz von T an, was als *T^3-Gesetz* von DEBYE bekannt ist.

Es sei noch darauf hingewiesen, daß sich dem monotonen Verlauf der Funktionen $c_v(T)$ bzw. $c_p(T)$ mehr oder weniger breite Spitzen oder Buckel überlagern, falls in dem festen Stoff *Phasenumwandlungen 2. Ordnung* erfolgen, und bei *Phasenumwandlungen 1. Ordnung* weisen die Funktionen sogar Sprünge auf (↑ 6.4.4.). Das Auffinden solcher Besonderheiten mit Hilfe der differentiellen Thermoanalyse, d. h. einer genauen Bestimmung des Temperaturgangs der spezifischen Wärmekapazität, bildet ein wichtiges Mittel bei der Suche nach solchen strukturellen Umwandlungen im Festkörper.

8.1.3.2. Bose-Einstein-Statistik. Die quantitative Theorie der spezifischen Wärmekapazität von Festkörpern geht davon aus, daß die Innere Energie U im wesentlichen von den quantisierten Gitterschwingungen oder *Phononen* bestimmt wird (↑ 8.1.2.). Beim Erwärmen werden die Gitterschwingungen zunehmend angefacht oder, anders ausgedrückt, es werden Phononen erzeugt, die sich über die Schwingungszustände der akustischen und optischen Zweige verteilen. Umgekehrt schlafen die Schwingungen beim Abkühlen nacheinander ein, wobei Phononen vernichtet werden. Nun kann jeder Schwingungszustand Energie nur in ganzen Vielfachen der Energie eines Phonons $\hbar\Omega_q$ aufnehmen oder abgeben, wobei Ω_q die Kreisfrequenz der Gitterschwingung mit dem Wellenzahlvektor $\boldsymbol{q}$ ist. In der Nähe des absoluten Nullpunkts ist die Energie, die im Mittel bei der atomaren Bewegung übertragen wird, jedoch so klein, daß sie nicht zur Erzeugung von Phononen ausreicht. Infolgedessen können die Gitterschwingungen überhaupt nicht angefacht werden, und es leuchtet ein, warum die spezifische Wärmekapazität der Festkörper für $T = 0\,\mathrm{K}$ verschwinden muß. Bei allmählichem Aufheizen kann mehr Energie zugeführt werden, so daß die Gitterschwingungen nacheinander erwachen, und zwar zunächst diejenigen der akustischen Phononenzweige mit der kleinsten Kreisfrequenz. Die Energie der optischen Phononen (s. a. Abb. 8.1.2-5) ist dagegen so hoch, daß diese Schwingungsmoden selbst bei Zimmertemperatur vielfach nur unvollständig an der Wärmebewegung teilnehmen können.

Um dieses qualitativ richtige Bild zu einer quantitativen Theorie auszubauen, muß man wegen der großen Zahl von Schwingungszuständen mit statistischen Methoden arbeiten. Man betrachtet die Gesamtheit der Phononen dazu als eine Art Gas von Quasiteilchen und fragt, wie sich die Phononen bei gegebener Temperatur auf die Schwingungszustände verteilen. Die richtige Aussage liefert die Bose-Einstein-Statistik, nach der Phononen als unterscheidbare Teilchen angesehen werden, von denen im Prinzip beliebig viele in einem Schwingungszustand untergebracht sein können. Diese Art der Statistik unterscheidet sich von der Boltzmannstatistik

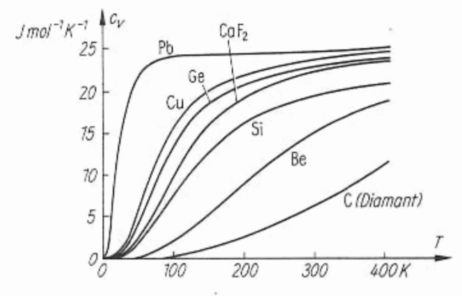


Abb. 8.1.3-2 Molare Wärmekapazitäten als Funktion der Temperatur für verschiedene Stoffe

klassischer Teilchen durch das Prinzip der Nichtunterscheidbarkeit aus der Quantentheorie (↑ 2.5.5.3.) und von der für Elektronen gültigen Fermi-Dirac-Statistik (↑ 8.2.2.2.) in der Besetzungsmöglichkeit der Zustände mit vielen Teilchen. Für die mittlere Anzahl von Phononen

$\langle n_q \rangle$, die sich bei der Temperatur T in einem Schwingungszustand Ω_q aufhalten, gilt nach der Bose-Einstein-Statistik folgende Verteilungsfunktion:

(26) $\langle n_q \rangle = 1/[\exp(\hbar\Omega_q/k_B T) - 1]$.

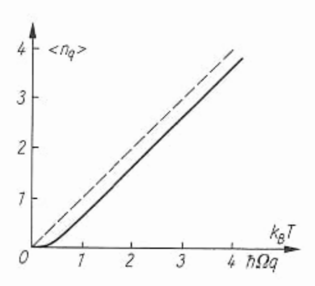


Abb. 8.1.3-3 Verlauf der Verteilungsfunktion von BOSE und EINSTEIN; bei großem T steigt $\langle n_q \rangle$ annähernd linear

Der Verlauf dieser Funktion ist in Abb. 8.1.3-3 dargestellt. Da $\langle n_q \rangle$ einen statistischen Mittelwert darstellt, was mit den spitzen Klammern unterstrichen werden soll, sind natürlich auch gebrochene Besetzungszahlen zulässig, und die Funktion verläuft kontinuierlich.

8.1.3.3. Einsteinnäherung. Ein erstes Modell zur quantitativen Beschreibung der molaren Wärmekapazität von Festkörpern wurde bereits 1907 von EINSTEIN aufgestellt. In drastischer, aber geschickter Vereinfachung setzte er die $3N$ Schwingungszustände eines Kristalls mit N Atomen als gleichartige Oszillatoren mit der einheitlichen Frequenz Ω an. Die innere Energie, die ein Mol des Festkörpers mit N_A Atomen (N_A Avogadrozahl) enthält, ergibt sich dann als Produkt des Ausdrucks 8.1.(26) mit $3N_A\hbar\Omega$ zu $U = 3N_A\hbar\Omega[\exp(\hbar\Omega/k_B T) - 1]^{-1}$. Durch Differenzieren nach T folgt für die molare Wärmekapazität die *Einsteinfunktion*

(27) $c_{vm} = 3R(\Theta/T)^2 \exp(\Theta/T)[\exp(\Theta/T - 1)]^{-2}$,

wobei die Abkürzung $\Theta \equiv \hbar\Omega/k_B$ eingeführt und von dem Zusammenhang $R = N_A k_B$ zwischen molarer Gaskonstante R und Boltzmannkonstante k_B Gebrauch gemacht wurde. Für den Grenzfall hoher Temperaturen folgt durch Entwickeln der Exponentialfunktionen gemäß $\exp x \cong 1 + x + \ldots$ und Abbruch nach dem 1. Glied $c_{vm} = 3R$ also der Wert entsprechend der Dulong-Petitschen Regel. Im umgekehrten Grenzfall sehr niedriger Temperaturen kann die 1 im Nenner von 8.1.(27) vernachlässigt werden, ein Exponentialglied läßt sich dann kürzen, und man sieht, daß c_{vm} mit Annäherung an den absoluten Nullpunkt exponentiell gegen Null geht. Mit dieser Deutung des qualitativen Verlaufs der molaren Wärmekapazität leistete EINSTEIN einen bedeutenden Beitrag sowohl für die sich entwickelnde Quantentheorie als auch für die entstehende Festkörperphysik. Darüber hinaus vermag die Einsteinfunktion den Gang von c_{vm} bei höheren Temperaturen auch quantitativ richtig zu beschreiben, wenn Ω passend bestimmt wird. Dies ist verständlich, weil die optischen Schwingungsmoden, die bei sinkender Temperatur zuerst einschlafen, nur ein relativ schmales Frequenzband überdecken und daher näherungsweise als Oszillatoren gleicher Frequenz beschrieben werden können (s. a. Abb. 8.1.2-4).

8.1.3.4. Debyenäherung. DEBYE und Mitarbeiter arbeiteten ab 1910 eine Theorie der molaren Wärmekapazitäten aus, bei der die Gitterschwingungen unter Vernachlässigung des atomaren Kristallbaus als Schallwellen in einem elastischen Kontinuum aufgefaßt werden. Diese Näherung ist im Unterschied zur Einsteinschen Behandlung vor allem bei tiefen Temperaturen gerechtfertigt, da in diesem Bereich vorwiegend Phononen mit kleiner Frequenz bzw. großer Wellenlänge zur inneren Energie beitragen. Zur Berechnung des Beitrags der einzelnen Schwingungsmoden wird das Produkt aus der Verteilungsfunktion [8.1.(26)], der Zustandsdichtefunktion [8.1.(27)] und der Phononenenergie $\hbar\Omega_q$ gebildet. Die Summen über die einzelnen Schwingungszustände werden durch Integrale ersetzt, was wegen der relativ dichten Folge von Schwingungsmoden gerechtfertigt ist. Eine obere Grenzfrequenz Ω_D bis zu der die Integration zu erstrecken ist, wird mit einem Kunstgriff so festgelegt, daß als Gesamtzahl der Schwingungszustände eines Kristalls aus N Atomen gerade $3N$ herauskommt. Abb. 8.1.3-4 veranschaulicht den Verlauf der Zustandsdichtefunktion für eine lineare Atomkette (s. a. 8.1.2.3.) im Vergleich zum schraffierten Rechteckverlauf der Debyenäherung. Im Ergebnis der etwas umständlichen

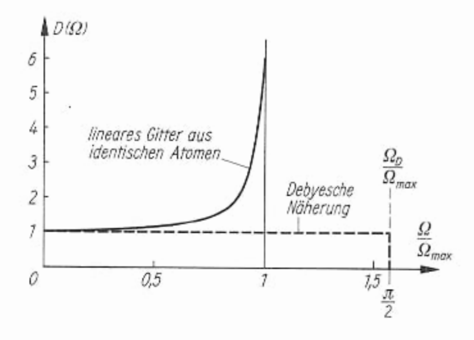


Abb. 8.1.3-4 Verlauf der Zustandsdichte in der Debeyeschen Näherung und einem Atomgitter im eindimensionalen Modell

Rechnung ergibt sich für die molare Wärmekapazität die *Debyefunktion*

$$(28)\quad c_{vm} = 9R(\Theta/T)^{-3} \int_0^{\Theta/T} [\exp(\Theta/T)] \times (\Theta/T)^4[\exp(\Theta/T - 1]^{-2}\, d(\Theta/T),$$

mit der sich der Verlauf von c_{vm} für viele Festkörper recht gut wiedergeben läßt. $\Theta \equiv \hbar\Omega_D/k_B$ ist die charakteristische oder *Debyetemperatur*, die vielfach zur Charakterisierung von Festkörpern benutzt wird. Die Grenzfrequenz Ω_D und damit auch die Debyetemperatur sind um so höher, je stärker die Bindungskräfte und je leichter die Atome im Festkörper sind. Zahlenwerte für einige Stoffe bringt Tab. 8.1.3-5.

Tab. 8.1.3-5 Debyetemperatur Θ in K für verschiedene feste Elemente

Stoff	Θ	Stoff	Θ
Li	344	Cu	343
C	2230	Zn	327
Na	158	Ge	374
Mg	400	Se	90
Al	428	Ag	225
Si	645	W	400
Fe	470	Pb	105
Ni	450	U	207

Auch aus der Debyefunktion folgt für hohe Temperaturen $T \gg \Theta$ der klassische Wert $c_{vm} = 3R$. Bei tiefen Temperaturen $T \ll \Theta$ läßt sich die Funktion durch das Debyesche T^3-Gesetz in der Form $c_{vm} = 3R(12\pi^4/15)(T/\Theta)^3$ annähern. Im Zwischenbereich sind für die analytisch nicht geschlossen auswertbaren Integrale der Debyefunktion Werte aus Tabellenwerken zu entnehmen. Speziell bei Kristallen von Verbindungen, bei denen auch die akustischen Phononen zur Wärmekapazität beitragen, ergeben sich im Bereich höherer Temperaturen vielfach stärkere Abweichungen zum Verlauf der Debyefunktion. Empirisch hilft man sich häufig, indem Anteile der Debye- und der Einsteinfunktion überlagert und die Konstanten an den gemessenen Verlauf angepaßt werden.

8.1.3.5. Gitterdynamische Theorie. Seit der Begründung der Quantentheorie der Gitterschwingungen durch BORN und VON KARMAN, die fast gleichzeitig zur Entwicklung des Debyeschen Kontinuummodells erfolgte, ist mit immer genaueren Rechnungen und zunehmend unter Einsatz von Computern versucht worden, eine konsequente atomistische Theorie des Verlaufs der molaren Wärmekapazität von Festkörpern zu entwickeln. Man muß dazu zunächst das Spektrum der Schwingungsmoden genau berechnen (s. a. 8.1.2.3. und 8.1.2.4.) und unter Benutzung der Bose-Einstein-Verteilungsfunktion [8.1.(26)] durch Aufsummieren die Beiträge zur inneren Energie ermitteln. Insbesondere für Metalle konnten bei Berücksichtigung mehrerer tausend einzelner Schwingungsmoden beachtliche Resultate erzielt werden. Dabei gelang es, auch Feinheiten zu deuten, die als *van-Hove-Singularitäten* der Zustandsdichte oder als *Kohnanomalien* in den Dispersionskurven der Gitterschwingungen bekannt geworden sind. Aus derartigen Besonderheiten kann beispielsweise auf die Besetzung der Elektronenzustände geschlossen werden (s. a. 8.4.3.4.).

Abschließend soll darauf hingewiesen werden, daß vor allem bei Metallen auch die Elektronen einen mit der Temperatur ansteigenden Beitrag zur *Wärmekapazität von Festkörpern* liefern (↑ 8.4.3.3.).

8.1.4. Anharmonische Effekte, Wärmeausdehnung, Wärmeleitung

8.1.4.1. Anharmonizität der Gitterschwingungen. Bei allen bisherigen Überlegungen wurde vorausgesetzt, daß die Gitterschwingungen harmonisch sind, also durch Wellen mit sinusförmigem Verlauf beschrieben werden können. Als strenge Lösungen werden *harmonische Wellen* erhalten, wenn die rücktreibende Kraft, die auf einen Gitterbaustein wirkt, proportional zur Auslenkung u von der Gleichgewichtsposition a_0 ist (s. a. 8.1.2.3.). In diesem Fall hängt die potentielle Energie wie beim harmonischen Oszillator quadratisch von der Auslenkung ab; die Potentialmulde besitzt in einer Dimension den Verlauf einer Parabel. Aus der Theorie der *Bindungskräfte im Festkörper* (↑ 6.4.3.) folgt jedoch ein asymmetrischer Verlauf, wie er in Abb. 8.1.4-1 dargestellt ist. Der linke Zweig, der einer Annäherung der Atomrümpfe entspricht, verläuft sehr steil, weil die inneren Elektronenhüllen

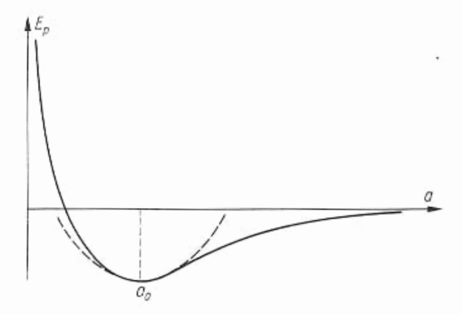


Abb. 8.1.4-1 Reales Wechselwirkungspotential als Funktion des Atomabstands und harmonische Näherung (*gestrichelt*); a Atomabstand, E_p potentielle Energie

einer Deformation großen Widerstand entgegensetzen. Der Verlauf des rechten Zweigs hängt von der Natur der Bindungskräfte ab. Er läßt sich durch Potenz- oder Exponentialausdrücke beschreiben, die vom Parabelverlauf ebenfalls erheblich abweichen. Die harmonische Näherung für die Gitterschwingungen gilt daher nur für den Bereich einigermaßen, der sich in der Nähe des Minimums der Potentialmulde durch eine Parabel annähern läßt. Daß die harmonische Näherung für die Beschreibung der meisten Prozesse, an denen Gitterschwingungen beteiligt sind, ausreichend erfüllt ist, ergibt sich aus der relativ kleinen Auslenkung der Gitterbausteine im Rahmen der thermischen Bewegung. Immerhin kann die maximale Verschiebung von der Gleichgewichtsposition bei Festkörpern dicht unterhalb des Schmelzpunkts nahezu 10% des mittleren Atomabstands erreichen.

Trotz der meist kleinen Abweichungen von der harmonischen Schwingungsform darf die *Anharmonizität der Gitterschwingungen* in der Festkörperphysik jedoch nicht übersehen werden, denn erst ihre Berücksichtigung liefert den Schlüssel für das qualitative Verständnis wichtiger Erscheinungen, wie in den folgenden Abschnitten anhand von Beispielen dargelegt wird. Eine quantitative Behandlung anharmonischer Gitterschwingungen führt auf nichtlineare Differentialgleichungen und damit zu großen mathematischen Schwierigkeiten, die erst neuerdings schrittweise überwunden werden. Die Phononen sind dann keine strengen Eigenlösungen der Gleichungen, während die den strengen Lösungen entsprechenden nichtlinearen Wellen im allgemeinen keine zeitlich stationäre Bewegungsform verkörpern. Eine Ausnahme stellen bestimmte singuläre Lösungen in Form von Wellenpaketen, sogenannte *Solitons*, dar. Ihre Bedeutung in verschiedenen Bereichen der Physik, darunter der Festkörperphysik, bildet gegenwärtig einen Schwerpunkt der Forschung. In diesem Zusammenhang sind auch stark nichtlineare Effekte zu erwähnen, die bei Stoßwellen in Festkörpern oder anderen Medien im Gefolge von Explosionen oder der Bestrahlung mit kräftigen Laserimpulsen auftreten. Solche *Stoßwellen* können in Form einer dynamischen Höchstdruckreaktion (↑ 9.3.9.) Phasenumwandlungen in Festkörpern bewirken. Ihre systematische Erforschung steht jedoch noch am Anfang.

8.1.4.2. Wärmeausdehnung der Festkörper. Nahezu alle Festkörper dehnen sich beim Erwärmen geringfügig aus. Für eine Probe der Länge l bzw. des Volumens V definiert man durch die relativen Änderungen beim Temperaturanstieg ΔT den *linearen* bzw. den *kubischen Ausdehnungskoeffizienten*:

(29) $\Delta l/l = \alpha_l \Delta T; \quad \Delta V/V = \alpha \Delta T; \quad \alpha \cong 3\alpha_l.$

Nach der letzten Beziehung beträgt der kubische oder Volumenausdehnungskoeffizient etwa das Dreifache des linearen. Werte des linearen Koeffizienten sind für einige Werkstoffe in Tab. 8.1.4-2 wiedergegeben. Trotz der relativen Kleinheit spielt die Wärmeausdehnung der Festkörper, die übrigens bei nichtkubischen Einkristallen anisotrop oder richtungsabhängig ist, in Natur und Technik eine wichtige Rolle. Die von ihr bewirkten Spannungen können Anlaß für Deformationen, Risse und Brüche sein, was bei der Konstruktion von Eisenbahngleisen, Hallen oder Anlagen sowie bei der Ausführung von Verbindungen durch Nieten oder Schweißen beachtet werden muß. Umgekehrt wird die Verbiegung von Bimetallstreifen, die aus zwei Legierungen mit unterschiedlichem Ausdehnungskoeffizienten zusammengesetzt sind, für die Auslösung von Schalt- und Regelsignalen genutzt. Die erste bewußte Anwendung der Wärmeausdehnung war wohl das bis zum Mittelalter übliche Sprengen von Gestein durch abwechselndes Erhitzen mit Holzfeuer und rasches Abschrecken mittels Wasser.

Atomistisch läßt sich die Wärmeausdehnung fester Stoffe aufgrund der Asymmetrie des anharmonischen Potentials von Abb. 8.1.4-1 leicht verstehen. Mit steigender Temperatur wird die Amplitude der Schwingungen um die Gleichgewichtsposition größer. Die Atome halten sich dann im zeitlichen Mittel zunehmend nach rechts verschoben, also weiter voneinander entfernt, auf. Bei rein harmonischen Schwingungen in einem Parabelpotential würde hingegen keine

Tab. 8.1.4-2 Lineare Wärmeausdehnungskoeffizienten α_l in $10^{-5}\,K^{-1}$ polykristalliner bzw. amorpher Werkstoffe

Stoff	α_l
Aluminium	2,2
Blei	2,7
Bronze	1,8
Chromium	0,8
Kupfer	1,6
Messing	1,9
Gußeisen	1,2
Stahl	0,9
Platin	0,9
Beton	0,75 bis 1,4
Fensterglas	1,0
Quarzglas	0,06
Porzellan	0,4
Granit	0,83
Kautschuk	7,7
PVC (hart)	7,5
Phenoplaste	2,2
Ziegelstein	0,3 bis 1,2

Ausdehnung resultieren, weil der mittlere Aufenthaltsort unabhängig von der Schwingungsamplitude gleich bliebe.

8.1.4.3. Wärmeleitung fester Stoffe. Werden die beiden Enden eines Stabs der Länge L und des Einheitsquerschnitts auf den konstanten Temperaturen T_1 bzw. T_2 gehalten, so fließt ein konstanter Wärmestrom $\Phi = -\lambda(T_2 - T_1)/L$, wobei $T_2 > T_1$ sein soll. Allgemein gilt, daß der Wärmestrom in differentiellen Bezirken mit dem gleichen Faktor λ dem Temperaturgradienten $\partial T/\partial r$ proportional ist. Die Stoffkennzahl λ, die bei Einkristallen von der Richtung abhängen kann, wird *Wärmeleitzahl* oder *Wärmeleitvermögen* genannt; der gleichfalls benutzte reziproke Wert heißt (spezifischer) *Wärmewiderstand*. Die internationale Einheit ist $\mathrm{Wm^{-1}K^{-1}}$, doch werden in Tabellen oft noch die Maße $\mathrm{Wcm^{-1}K^{-1}}$ oder $\mathrm{cal\,cm^{-1}s^{-1}K^{-1}}$ benutzt. Für eine Reihe fester Materialien sind λ-Werte in Tab. 8.1.4-3 angegeben. Auch diese Größe ist technisch von sehr großer Bedeutung, beispielsweise in der Bautechnik, aber auch in der Elektrotechnik und Elektronik, da die Wärmeisolation im Hinblick auf möglichst geringe Energieverluste, andererseits jedoch eine optimale Wärmeabfuhr, etwa bei mikroelektronischen Schaltungen, wesentlich von der Auswahl passender Werkstoffe bestimmt wird. Besonders kleine Wärmeleitzahlen weisen amorphe Stoffe wie Gläser auf. Die extrem kleinen Werte von porösen Baustoffen sind allerdings weniger eine Festkörpereigenschaft, sondern eine Folge der eingeschlossenen Gasbläschen (s. a. 6.2.5.). Hohe Werte findet man bei Metallen sowie bei einigen Isolatorkristallen wie Saphir. Für die Kryotechnik ist wichtig, daß das Wärmeleitvermögen besonders bei niedrigen Temperaturen stark von der Temperatur abhängt (Abb. 8.1.4-4).

Tab. 8.1.4-3 Wärmeleitzahlen λ in $10^2\,\mathrm{Wm^{-1}K^{-1}}$ einiger Werkstoffe bei 20 °C

Stoff	λ
Kupfer	3,94
Silber	4,19
Aluminium	2,30
Messing	1,13
Reineisen	0,75
C-Stahl	0,46
Cr–Ni-Stahl	0,13
Blei	0,33
Beton	0,006 bis 0,011
Ziegelstein	0,005
Holz	0,0012 bis 0,0018
Stahlbeton	0,013
Papier	0,001 bis 0,002
Glaswolle	0,00031
Asbest	0,0007
Schaumplast	0,0002 bis 0,0003

Bei der atomistischen Deutung der Wärmeleitung in Festkörpern ist davon auszugehen, daß der Energiestrom von Phononen oder von Elektronen getragen sein kann. Der *Elektronenanteil*, der bei Metallen vorherrschend ist, wird in 8.4.2.2. betrachtet. Hier begnügen wir uns mit einer kurzen Diskussion des *Phononenanteils*, der bei elektrisch isolierenden Stoffen maßgeblich ist. Nach einem von DEBYE entwickelten Modell werden die Ansätze für den *Wärmetransport in idealen Gasen* (↑ 6.2.5.) auf das Phononengas des Festkörpers übertragen. Man findet dann für die Wärmeleitzahl $\lambda = 1/3\; v_s C_v L$, wo v_s die Schallgeschwindigkeit, C_v die Wärmekapazität und L die mittlere freie Weglänge der Phononen ist. Die letzte Größe wird von der Phononenstreuung an Oberflächen und Gitterdefekten, vor allem aber durch die *Phonon-Phonon-Streuung* bestimmt. Bei rein harmonischen Schwingungen dürften Phononen freilich gar nicht aneinander gestreut werden. Erst die geringe Anharmonizität führt dazu, daß ein sich ausbreitendes Phonon das Gitter unter dem Einfluß eines anderen Phonons aperiodisch verzerrt „sieht", wodurch es zur Streuung und zum Energieaustausch kommt. Analog zur *Photon-Phonon-Streuung* [8.1.(25)] gilt für die Phonon-Phonon-Streuung

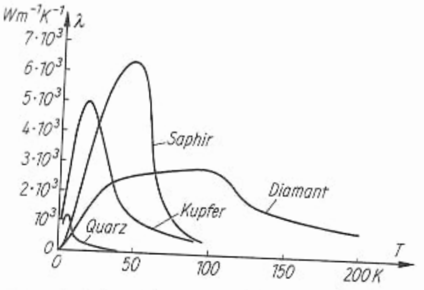


Abb. 8.1.4-4 Wärmeleitzahlen einiger Isolatoreinkristalle und von Kupfer als Funktion von T

$$(30)\quad \boldsymbol{q}_3 = \boldsymbol{q}_1 + \boldsymbol{q}_2 \pm \boldsymbol{G},$$

wobei die Indizes drei verschiedene Phononen und $\boldsymbol{G}$ einen reziproken Gittervektor kennzeichnen. Die Mitwirkung von elastischen Reflexionen am Gitter ist, wie PEIERLS zeigte, sogar unumgänglich, wenn sich die Phononen mit der Umgebung ins Wärmegleichgewicht setzen sollen. Ein wesentlicher Unterschied zum Wärmetransport in Gasen besteht auch darin, daß die Phononenzahl nicht konstant ist, sondern mit sinkender Temperatur abnimmt. Am heißen Ende eines Stabs werden fortwährend neue Phononen geschaffen, die Energie zum kalten

Ende transportieren, wobei sie längs des Wegs teilweise vernichtet werden.

8.1.4.4. Zustandsgleichung fester Stoffe. Ähnlich zu den Gasen ist häufig versucht worden, insbesondere von GRÜNEISEN, allgemeine Beziehungen zur Beschreibung mechanischer und thermischer Zustandsgrößen von Festkörpern aufzustellen. Obwohl diesen Bemühungen kein genereller Erfolg beschieden war, was angesichts der Spezifik des atomaren Baus der Festkörper nicht verwundern kann, sind im Ergebnis doch einige nützliche Näherungsformeln aufgestellt worden. Erwähnt sei die *Grüneisenbeziehung* $\alpha B V = \Gamma c_v$, die den Volumenausdehnungskoeffizienten α, den Kompressionsmodul B (das Reziproke der Kompressibilität χ), das Molvolumen V und die molare Wärmekapazität c_v verknüpft. Der Stoffparameter Γ liegt in der Nähe von 2 und heißt *Grüneisenkonstante*.

8.2. Bändermodell der Kristallelektronen

8.2.1. Elektronenzustände im Festkörper

In den folgenden Abschnitten wird das Verhalten der Elektronen im Festkörper behandelt. Die energetische Verteilung und die Bewegung der Elektronen in Festkörpern bilden nicht nur die Grundlage für das Verständnis beim Einsatz fester Stoffe in der Elektrotechnik und Elektronik, sondern auch die optischen und magnetischen Eigenschaften werden maßgeblich von der elektronischen Struktur bestimmt.

Die *elektrische Leitfähigkeit* ist die augenfälligste Materialeigenschaft, in der die recht unterschiedlichen elektronischen Eigenschaften der Festkörper zum Ausdruck kommen. Sie wird quantitativ durch den *spezifischen elektrischen Widerstand* ϱ oder durch die reziproke Größe, die *spezifische Leitfähigkeit* $\sigma = 1/\varrho$, angegeben. Da ein Leiter der Länge l und des Querschnitts A nach dem Ohmschen Gesetz den Widerstand $R = \varrho l/A$ besitzt, ist ϱ der Widerstand eines Würfels der Kantenlänge 1 m. Die Dimension von ϱ ist Ωm und $(\Omega\text{m})^{-1}$ für σ. Normalerweise kommt die elektrische Leitfähigkeit fester Stoffe durch den Transport von Elektronen unter dem Einfluß elektrischer Felder zustande. Neben dieser *Elektronenleitung* ist allerdings bei manchen Stoffen, insbesondere bei Ionenkristallen und Gläsern, eine *Ionenleitung* zu berücksichtigen, die hier ausgeklammert wird.

Nach ihrer unterschiedlichen elektrischen Leitfähigkeit klassifiziert man die festen Werkstoffe häufig in *Metalle* (oder metallische Leiter), *Halbleiter* und *Isolierstoffe*. Sehr gute Isolatoren können einen spezifischen Widerstand von mehr als 10^{20} Ωm aufweisen, Halbleiter besitzen bei Zimmertemperatur ϱ-Werte im Bereich von 10^7 bis herab zu 10^{-4} Ωm, während für gute metallische Leiter bei tiefen Temperaturen Werte von 10^{-9} bis 10^{-10} Ωm gefunden werden. Schließlich kann die Leitfähigkeit bei Supraleitern (↑ 8.5.) sogar auf unmeßbar kleine Werte, also praktisch auf Null, absinken.

Jeder Versuch, das Verhalten der Elektronen im Festkörper zu beschreiben, muß letztlich daran gemessen werden, ob die über viele Größenordnungen variierende elektrische Leitfähigkeit im Rahmen eines atomistischen Modells gedeutet werden kann. Weder die Auffassung der Leitungselektronen als eine Art Gas aus klassischen Teilchen, noch die Vorstellung von Elektronen, die in den zwischenatomaren Bindungen lokalisiert sind, führten zu einem befriedigenden Ergebnis. Den Schlüssel für ein Verständnis lieferte erst die Anwendung der Quantentheorie auf den Festkörper als Gesamtsystem. Allerdings würde die strenge Lösung des quantenmechanischen Vielkörperproblems für eine sehr große Anzahl von Atomkernen und Elektronen auf unüberwindliche Schwierigkeiten stoßen. Man hilft sich daher mit folgenden Vereinfachungen:

Es wird vorausgesetzt, daß die Elektronen abgeschlossener Schalen durch den Zusammenschluß der Atome im Verband des Festkörpers nur unwesentlich beeinflußt werden, also lokalisiert an den Atomrümpfen verbleiben. Dagegen sollen die Außen- oder Bindungselektronen dem Kristall als Ganzes angehören und sich in Form von Materiewellen ausbreiten können.

Man untersucht nun Lösungen der *Schrödingergleichung* (↑ 2.5.2.2.) für ein einzelnes, als repräsentativ herausgegriffenes Elektron, das sich im räumlich-periodischen Potential der Atomrümpfe eines als ideal angenommenen Kristalls bewegt. Die Wechselwirkung mit den übrigen Außenelektronen wird im Rahmen dieser *Einelektronennäherung* nur pauschal durch einen Potentialterm berücksichtigt.

Praktisch bereitet die Berechnung der Energiezustände und der Eigenfunktionen für diese *Kristallelektronen*, wie man die Außenelektronen gewöhnlich nennt, immer noch große mathematische Schwierigkeiten. Mit Hilfe verschiedener Näherungsverfahren ist es jedoch gelungen, ein immer genaueres Bild vom Verhalten der Elektronen im Festkörper zu entwerfen, das als *Bändermodell der Kristallelektronen* bekannt geworden ist.

8.2.1.1. Näherung freier Elektronen. Einen Zugang zur Begriffsbildung des Bändermodells vermittelt die Näherung freier Elektronen, die zuerst von SOMMERFELD eingeführt wurde. Bei diesem drastisch vereinfachten Modell vernachlässigt man den periodischen Verlauf des Potentials im Innern des Kristalls und nimmt dafür einen konstanten mittleren Wert des Potentials an. Da die potentielle Energie stets nur bis auf

eine Konstante definiert ist, kann man $V = 0$ setzen. Die Kristalloberflächen werden durch senkrechte Potentialwände wiedergegeben, die bis zu dem konstanten Außenpotential V_a reichen. Damit ergibt sich ein rechteckiger Potentialkasten, der in einer Dimension in Abb. 8.2.1-1 dargestellt ist. Als Lösung der Schrödingergleichung für das Innere des Kastens er-

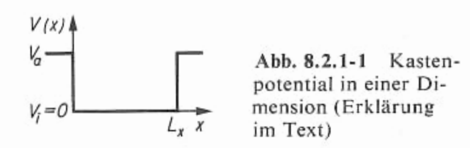


Abb. 8.2.1-1 Kastenpotential in einer Dimension (Erklärung im Text)

geben sich ebene Wellen (↑ 2.5.4.3.), die in komplexer Schreibweise durch

(1) $\psi_k = c \exp \mathrm{i}(kx \pm \omega t)$
$= c \exp (\mathrm{i}kx) \exp \pm (\mathrm{i}\omega t)$

beschrieben werden. Das sind Eigenfunktionen freier Elektronen, die sich als fortlaufende Materiewellen mit der Wellenzahl $k = 2\pi/\lambda$ und der Kreisfrequenz $\omega = E_k/\hbar$ (E_k Energie des Elektrons) ausbreiten. Üblicherweise wird der zeitabhängige Anteil $\exp \pm (\mathrm{i}\omega t)$ weggelassen, da er für die Betrachtung stationärer Lösungen unwesentlich ist. Die Konstante c ist durch Integration von $\psi_k\psi_k^*$ über den Aufenthaltsraum entsprechend der Normierungsforderung der Quantentheorie so festzulegen, daß die Gesamtwahrscheinlichkeit, das Elektron anzutreffen, gleich Eins wird.

Nun sind die Elektronen unseres fiktiven Kristalls jedoch nicht völlig frei, sondern im Potentialkasten eingesperrt. Die Auswahl der stationären Elektronenzustände aus dem Kontinuum der ψ_k wird mit Hilfe geeigneter Randbedingungen vorgenommen. Physikalisch sinnvolle Ergebnisse liefern *periodische Randbedingungen* der Form $\psi_k(x + L) = \psi_k(x)$, wonach die Wellenfunktionen an den Kastenwänden bei $x = 0$ und $x = L$ die gleiche Phase haben müssen. Damit ergibt sich eine weitgehende Analogie zur Abzählung der Phononenzustände, und die erlaubten Wellenzahlen betragen $k = 2\pi n/L$ mit $n = 0, \pm 1, \pm 2 \ldots$ (↑ 8.1.2.3.). Erweitern wir diese Überlegungen auf einen dreidimensionalen Potentialtopf mit der Kantenlänge L und dem Volumen $V = L^3$, so sind die periodischen Randbedingungen auf ebene Wellen im Raum mit dem Wellenzahlvektor $\boldsymbol{k}$ anzuwenden. Für den Betrag der erlaubten $\boldsymbol{k}$-Werte folgt $\boldsymbol{k}^2 = k_x^2 + k_y^2 + k_z^2 = (4\pi^2/L^2)(n_x^2 + n_y^2 + n_z^2)$, wobei k_x, k_y und k_z die räumlichen Komponenten des Wellenzahlvektors sind, während die ganzen Zahlen n_x, n_y und n_z als Quantenzahlen für die Numerierung der Elektronenzustände im Festkörper angesehen werden können. Einen anschaulichen Überblick über die Verteilung dieser Zustände gewinnt man durch ihre Darstellung im $\boldsymbol{k}$-Raum mit den Koordinatenachsen k_x, k_y, k_z (↑ 8.1.2.4.). Beginnend mit $n_x = n_y = n_z = 0$ entspricht dann jedem Punkt, der zu einem ganzzahligen Tripel dieser Quantenzahlen gehört, ein erlaubter $\boldsymbol{k}$-Zustand. Die Zustandsdichte im $\boldsymbol{k}$-Raum ist demnach konstant, und sie beträgt entsprechend zu [8.1.(22)]

(2) $Z(\boldsymbol{k}) = (L/2\pi)^3 = V/8\pi^3$.

Allerdings ist die Anzahl dieser $\boldsymbol{k}$-Zustände der Elektronen im Unterschied zu den $\boldsymbol{q}$-Zuständen der Phononen nicht begrenzt, sondern es sind im Prinzip beliebig große $\boldsymbol{k}$-Werte möglich. Dagegen brauchen keine unterschiedlichen Polarisationen der Wellen berücksichtigt zu werden, da ja die ψ_k-Funktionen keine Schwingungsbewegung einer physikalischen Größe verkörpern (↑ 2.5.1.5.).

Im weiteren soll untersucht werden, wie die $\boldsymbol{k}$-Zustände mit den Eigenfunktionen ψ_k gemäß [8.2.(1)] mit Elektronen besetzt werden können und wie sich die so beschriebenen Kristallelektronen verhalten. Von großer Tragweite ist dabei die Tatsache, daß Elektronen als Teilchen mit dem Spin $\hbar/2$ dem *Pauliprinzip* (↑ 2.5.5.2.) unterliegen; infolgedessen kann jeder $\boldsymbol{k}$-Zustand höchstens zwei Elektronen, und zwar mit entgegengesetzter Spinorientierung, aufnehmen. Außerdem gilt für Elektronen als „realen" Teilchen zum Unterschied von den Phononen natürlich das Gesetz von der Erhaltung der Teilchenzahl, so daß im Festkörper bei den üblichen Energien, die weit unter der Schwelle für die Bildung von *Elektron-Positron-Paaren* (↑ 3.4.2.) liegen, Elektronen weder erzeugt noch vernichtet werden können.

Die Anzahl der Elektronenzustände im Festkörper ist also doppelt so groß wie die Anzahl

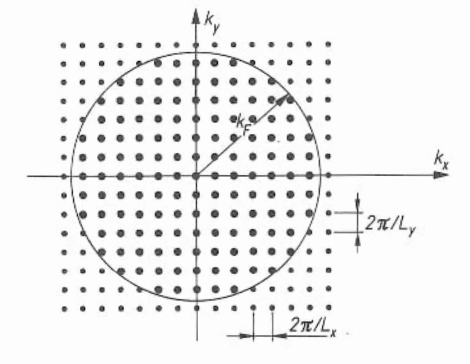


Abb. 8.2.1-2 Elektronenzustände im $\boldsymbol{k}$-Raum (zweidimensionaler Schnitt); bei $T = 0$ K sind alle Zustände innerhalb der Fermikugel mit dem Radius $|\boldsymbol{k}_F|$ besetzt

der k-Zustände, deren Dichte durch [8.2.(2)] gegeben ist. Auf diesen Zuständen sind unter Beachtung des Pauliprinzips die Elektronen unterzubringen, die von den Atomen des Festkörpers als Bindungselektronen eingebracht werden.
In der Darstellung des k-Raums entspricht jedem doppelt besetzbaren Zustand ein Punkt mit dem Wellenzahlvektor $\boldsymbol{k}$, was Abb. 8.2.1-2 für einen zweidimensionalen Schnitt veranschaulicht. Der Impuls eines freien Elektrons, das sich in einem solchen Zustand aufhält, beträgt aber nach der *de-Broglieschen Beziehung* (↑ 1.2.3.) $\boldsymbol{p} = \hbar \boldsymbol{k}$, d. h., zu jedem Zustand gehört ein nach Betrag und Richtung definierter *Impuls des Kristallelektrons*. Da der Impuls bestimmt ist, kann über den Aufenthaltsort des Elektrons nach der Heisenbergschen Unschärferelation natürlich keine Aussage getroffen werden. Tatsächlich erhält man für die Ortsverteilung, die durch das Produkt $\psi_k \psi_k^*$ beschrieben wird, nach [8.2.(1)] den konstanten Wert c^2. Im Rahmen der Näherung freier Elektronen werden die Kristallelektronen somit als Teilchen mit bekanntem Impuls, aber einer über den ganzen Kristall „verschmierten" Aufenthaltswahrscheinlichkeit betrachtet. Diese zunächst ungewöhnlich erscheinende Darstellung bietet den Vorteil, daß auch die besonders interessierende Energie der Kristallelektronen genau bestimmt ist. Der den Kristall repräsentierende Potentialtopf wird dazu als so groß angenommen, daß die Lokalisierung infolge der Kastenwände vernachlässigt werden kann.
Aus der klassischen Beziehung $E_k = p^2/2m$ zwischen kinetischer Energie, Impuls und Masse folgt durch Einsetzen von $\boldsymbol{p} = \hbar \boldsymbol{k}$ der für Kristallelektronen wichtige Zusammenhang

(3) $E_k = \hbar k^2/2m_e$ (m_e Elektronenmasse).

Quantentheoretisch gelangt man zum gleichen Ergebnis, indem die Energieeigenwerte zu den Funktionen ψ_k durch Einsetzen von (8.2.(1)) in die Schrödingergleichung bestimmt werden. Der Verlauf $E(\boldsymbol{k})$ weist daher die Gestalt einer Parabel auf, die allerdings genaugenommen nur für die Punkte definiert ist, die den erlaubten k-Werten entsprechen (Abb. 8.2.1-3).

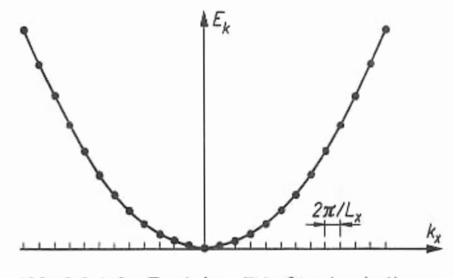


Abb. 8.2.1-3 Funktion $E(\boldsymbol{k})$ für ein eindimensionales Kastenpotential

Die weiterführende physikalische Interpretation der hier diskutierten Eigenschaften von *Elektronen im Potentialkasten* führt zu Ergebnissen, die für die Festkörperphysik außerordentlich bedeutsam sind. Infolge der Quantelung, die aus den Randbedingungen resultiert, können sich die Elektronen im Kristall nur in diskreten Zuständen bezüglich der Energie und des Impulses aufhalten. Da diese Zustände wegen der im Vergleich zu den atomaren Abständen großen Dimensionen eines Kristalls sehr zahlreich und eng benachbart sind, spricht man auch von einer *quasikontinuierlichen Verteilung der Zustände*, und häufig werden Energie und Wellenzahl der Kristallelektronen näherungsweise als kontinuierliche Variable aufgefaßt. Die diskrete Natur der Zustände und die durch das Pauliprinzip geregelte Besetzungsmöglichkeit haben jedoch zur Folge, daß sich die Kristallelektronen auch beim absoluten Nullpunkt der Temperatur über einen relativ breiten Energie- und Impulsbereich verteilen. Zur Wellenzahl 0, die ruhenden Elektronen entspricht, gehört ja nur ein k-Zustand, in dem lediglich zwei Elektronen Platz finden. Alle übrigen Kristallelektronen müssen sich in Zuständen mit $|\boldsymbol{k}| > 0$ aufhalten, und ihre kinetische Energie nimmt mit wachsendem $|\boldsymbol{k}|$ quadratisch zu. Bei $T = 0\,\mathrm{K}$ werden die Elektronen natürlich die energetisch tiefsten Zustände einnehmen. Am absoluten Nullpunkt sind daher alle Elektronenzustände bis zu der sogenannten Fermienergie E_F bzw. der Wellenzahl $|\boldsymbol{k}_F|$ vollständig besetzt, alle Zustände oberhalb E_F dagegen leer. In drei Dimensionen füllen die besetzten Zustände im k-Raum eine Kugel mit dem Radius $|\boldsymbol{k}_F|$ aus, die man als *Fermikugel* bezeichnet. Die Elektronen, die sich in den Zuständen an der Oberfläche dieser Kugel aufhalten, haben die Energie E_F. Ist die Anzahl der Kristallelektronen pro Volumeneinheit des wirklichen Raums N_e, so findet man für die Fermienergie durch Berechnung des Kugelvolumens mit dem Radius $|\boldsymbol{k}_F|$ im k-Raum unter Benutzung der Zustandsdichte [8.2.(2)]

(4) $E_F = [\hbar^2(3\pi^2 N_e)^{2/3}]/2m_e$.

Die praktischen Werte betragen für Metalle einige Elektronenvolt (↑ 8.4.3.4.). Heizt man den Festkörper auf, so wird die scharfe Grenze zwischen besetzten und unbesetzten Zuständen, d. h. die Oberfläche der Fermikugel „verwischt", indem Elektronen aus einem energetischen Gebiet etwa der Breite von $k_B T$ (k_B Boltzmannkonstante) durch die Wärmebewegung in unbesetzte Zustände oberhalb der Fermienergie angeregt werden. Für Temperaturen unterhalb des Schmelzpunkts der meisten Festkörper ist $k_B T$ stets viel kleiner als E_F, so daß auch bei höheren

Temperaturen die Mehrzahl der Kristallelektronen in den niedrigsten $\boldsymbol{k}$-Zuständen verharrt, weil mit der verfügbaren Energie keine unbesetzten Zustände erreicht werden können. Infolge dieses ungewöhnlichen Verhaltens, das sehr erheblich von dem eines klassischen Teilchensystems abweicht, wird das Elektronengas im Festkörper als *entartet* bezeichnet. Obwohl die *Näherung freier Elektronen*, auch *Modell freier Elektronen* genannt, nur einen Ausgangspunkt für die theoretische Behandlung der Kristallelektronen bildet, lassen sich damit bereits wichtige Eigenschaften der Leitungselektronen von Metallen beschreiben (↑ 8.4.3.). Ferner kann die wichtige Beziehung 8.2.(3) für das allgemeinere Bändermodell übernommen werden, wenn man die Elektronenmasse m_e durch eine *effektive Masse* m^* ersetzt. Das gilt auch für die *Zustandsdichtefunktion* $D(E)$, welche die Anzahl der Elektronenzustände pro Volumeneinheit als Funktion der Energie angibt. Man erhält sie, indem man die Dichte der $\boldsymbol{k}$-Zustände [8.2.(2)] mit dem Volumen einer Kugelschale im $\boldsymbol{k}$-Raum multipliziert und wegen der doppelten Besetzungsmöglichkeit einen Faktor 2 hinzufügt. Drückt man dann noch $\boldsymbol{k}$ mittels der Beziehung 8.2.(3) durch E aus, so ergibt sich

(5) $$D(E) = 1/2\pi^2(2m_e/\hbar^2)^{3/2}\,E^{1/2}.$$

Der Verlauf dieser Funktion ist in Abb. 8.2.1-4 dargestellt. Der schraffierte Bereich gibt an, welche Zustände bei $T = 0$ K besetzt sind.

8.2.1.2. Elektronenzustände im periodischen Potential

8.2.1.2.1. *Brillouinsche Näherung.* Entsprechend dem von BRILLOUIN begründeten Vorgehen wird zunächst untersucht, wie sich die Elektronenzustände der Näherung freier Elektronen verändern, wenn im Inneren des Potentialkastens eine schwache Störung durch einen periodischen Potentialverlauf berücksichtigt wird. An den Stellen der im Abstand a angeordneten Atomrümpfe weist die potentielle Energie tiefe Mulden auf, während dazwischen das konstante Potential des Kastenmodells erhalten bleiben soll (Abb. 8.2.1-5a). In nullter Näherung wird angenommen, daß die ebenen Elektronenwellen ψ_k nach [8.2.(1)] auch für dieses Problem geeignete Lösungen darstellen.

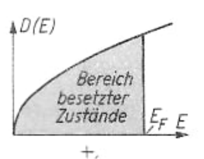


Abb. 8.2.1-4 Verlauf der Zustandsdichtefunktion $D(E)$, E_F Fermienergie

Eine wesentliche Modifizierung der Elektronenzustände resultiert nun daraus, daß die Potentialmulden als Streuzentren wirken können, an denen die Elektronenwellen abgebeugt bzw. reflektiert werden. Die Bedingungen hierfür entsprechen völlig der *Beugung von Elektronen- oder Röntgenstrahlen*, die von außen in einen Kristall eingestrahlt werden (↑ 6.4.2.2., 8.1.2.5.). Danach treten Beugungsmaxima durch konstruktive Überlagerung der an den einzelnen Atomrümpfen abgebeugten Teilwellen auf, wenn die Bragggleichung erfüllt ist. In der Betrachtungsweise des reziproken Gitters ist dies der Fall, falls die $\boldsymbol{k}$-Werte die Grenzen einer Brillouinzone erreichen. In einer Dimension liegen diese Grenzen bei ganzzahligen Vielfachen von π/a, so daß Beugungsmaxima für $k_x = n\pi/a$ ($n = 0, \pm 1, \pm 2$ ...) zu erwarten sind. Wie aus Rechnungen hervorgeht, liegen die Wellenzahlen der Kristallelektronen tatsächlich in einem Bereich, der diese *kritischen k-Werte* enthält; die Wellenlängen liegen in der Größenordnung der Atomabstände a.

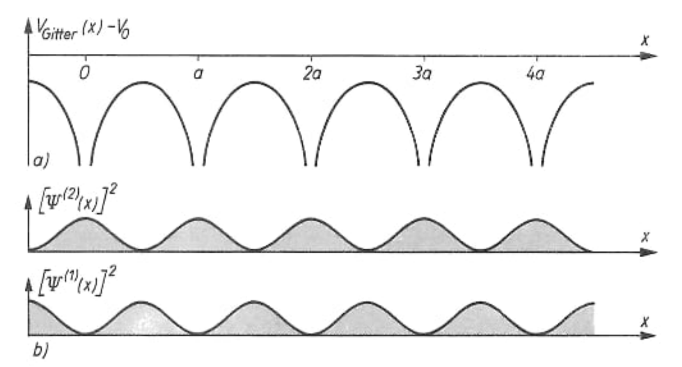


Abb. 8.2.1-5 a) Verlauf des periodischen Gitterpotentials, b) Aufenthaltswahrscheinlichkeit für Elektronen in Zuständen an den Zonengrenzen; a Atomabstand

Welche Auswirkungen ergeben sich nun für die Energiezustände der Kristallelektronen? Solange die Wellenzahlen von den kritischen Werten $\pm n\pi/a$ wesentlich verschieden sind, spielt die Beugung der Elektronenwellen keine erhebliche Rolle, und der parabelförmige Verlauf der Funktion $E(k)$ des Potentialkastenmodells wird praktisch nicht verändert. Mit Annäherung an

die Grenzen $k = \pm\pi/a$ der 1. Brillouinzone werden den Elektronenwellen in zunehmendem Maße Anteile einer reflektierten Welle beigemengt, was stärkere Abweichungen zur Folge hat. An den Zonengrenzen selbst überlagert sich der Elektronenwelle schließlich eine gleich starke reflektierte Welle, und es bauen sich stehende Wellen auf. Je nachdem, ob die Wellenfunktionen der Teilwellen addiert oder subtrahiert werden – beides ist bei der Bildung von Eigenfunktionen durch Linearkombination zulässig – ergeben sich zwei verschiedene stehende Wellen, nämlich

$$\psi^{(1)} = c[\exp(\mathrm{i}\pi x/a) - \exp(-\mathrm{i}\pi x/a)] = 2c \sin(\mathrm{i}\pi x/a) \text{ und}$$

$$\psi^{(2)} = c[\exp(\mathrm{i}\pi x/a) + \exp(-\mathrm{i}\pi x/a)] = 2c \cos(\mathrm{i}\pi x/a).$$

Trägt man die Aufenthaltswahrscheinlichkeiten ψ^2 dieser reellen Lösungen relativ zu den Positionen der Atomrümpfe auf (Abb. 8.2.1-5b), so findet man für eine dieser Lösungen Maxima an den Orten der Potentialmulden und Nullstellen dazwischen, während für die andere Lösung das Umgekehrte gilt. Im Vergleich zur konstanten Aufenthaltswahrscheinlichkeit der fortlaufenden Elektronenwellen des Potentialkastens entspricht der erste Fall einer Absenkung der Energie, weil sich das Kristallelektron vorzugsweise im Anziehungsbereich der Atomrümpfe aufhält, der zweite Fall dagegen einer Anhebung des Zustands. Im Energiebereich zwischen diesen beiden Zuständen existieren für den ungestörten Kristall keine stationären Lösungen. Durch den Einfluß des periodischen Potentials hat sich eine verbotene Zone in der Skala erlaubter Energiezustände ausgebildet, innerhalb der ein längerer Aufenthalt von Kristallelektronen nicht möglich ist. Die Elektronenzustände bis zu dieser Energielücke bilden das 1. Energieband. Da die zugehörige 1. Brillouinzone bei einem aus N Atomen bestehenden Kristall N erlaubte k-Werte umfaßt, lassen sich in dem Band maximal $2N$ Elektronen unterbringen.

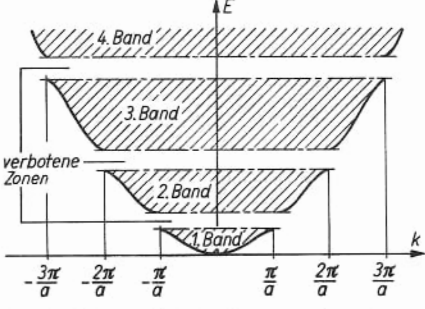


Abb. 8.2.1-6 $E(k)$-Funktion; Bänder und verbotene Zonen für ein schwaches periodisches Potential; a Atomabstand

Mit dem Eintritt in die 2. Brillouinzone, die sich von $k = +\pi/a$ bis $k = +2\pi/a$ und von $k = -\pi/a$ bis $k = -2\pi/a$ erstreckt, wiederholt sich der Verlauf im Prinzip. Zunächst weicht die Funktion $E(k)$ wegen des noch vorhandenen Einflusses der abgebeugten Teilwellen von der Parabel ab, danach ergibt sich im Inneren der Zone Übereinstimmung, schließlich kommt es bei der Annäherung an den zweiten kritischen k-Wert wieder zu Abweichungen und zur Ausbildung einer weiteren verbotenen Zone (Abb. 8.2.1-6). Diese Überlegungen lassen sich fortführen, jedoch interessieren meist nur zwei Energiebänder, und zwar das oberste bei $T = 0$ K gerade noch ganz oder teilweise gefüllte Band und das darüberliegende leere Band. Hier sei schon vorweggenommen, daß der Grad der Besetzung des einen Bands mit Elektronen und der Abstand zum darüberliegenden, beim absoluten Nullpunkt leeren Band, den Schlüssel für das Verständnis der Festkörperklassifikation in *metallische Leiter*, *Halbleiter* und *Isolatoren* bildet (↑ 8.2.3.1.).

8.2.1.2.2. *Blochsche Näherung.* Im vorangegangenen Abschnitt wurde von der Näherung freier Elektronen im Potentialkasten ausgegangen und der Einfluß einer schwachen Störung in Gestalt eines periodischen Potentials untersucht. Obwohl im Ergebnis eine qualitative Begründung für das Bändermodell der Kristallelektronen gegeben werden konnte, ist dieses Herangehen für stark gebundene Elektronen wenig realistisch. Die alternative Betrachtungsweise von Bloch wählt die Elektronenzustände und Eigenfunktionen der isolierten Atome als Ausgangspunkt, und es wird nach den Änderungen gefragt, die sich beim Zusammenschluß der Atome zum Festkörper ergeben. Für den einfachsten Fall der Annäherung von zwei gleichartigen Atomen ist das Ergebnis aus der *Quantentheorie der kovalenten chemischen Bindung* gut bekannt (s. a. 5.1.2.). Die Überlappung der ψ-Funktionen führt zu einer Aufspaltung ursprünglich gleicher, d. h. „entarteter" Eigenwerte der Atomelektronen in zwei neue Zustände, von denen der eine bindend und der andere nichtbindend ist. Quantenmechanische Rechnungen für Systeme aus N Atomen lehren, daß eine N-fache *Eigenwertaufspaltung* eintritt, wenn sich die Aufenthaltswahrscheinlichkeiten der Elektronen überlappen. Wir betrachten wieder eine eindimensionale Atomkette und lassen den Atomabstand a von großen Werten her abnehmen (Abb. 8.2.1-7). Solange der Abstand groß ist, ist die Kopplung zwischen den Atomzuständen nur gering, und es resultiert nur eine geringe Verbreiterung der Energieeigenwerte. Mit zunehmender Annäherung ergeben sich jedoch immer breitere Energiebänder, die verbotenen Zonen zwischen ihnen werden schmaler, schließlich kann es sogar zu einer Überlap-

pung der aus verschiedenen Atomzuständen hervorgehenden Energiebänder kommen. Ob der letztere Fall eintritt, hängt von der Energiedifferenz zwischen den Atomzuständen, der Natur ihrer Eigenfunktionen und vom Gleichgewichtsabstand der Atome a_0 ab. Wie die Blochsche Betrachtungsweise zeigt, spaltet im

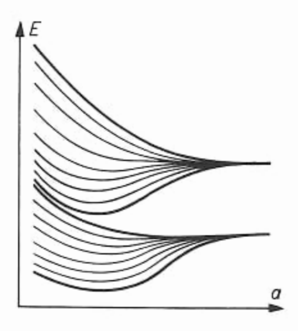


Abb. 8.2.1-7 Blochsche Näherung; Aufspaltung diskreter Energiezustände in Bänder mit je N Zuständen, die sich bei kleinem Atomabstand a teilweise überlappen können

Prinzip jeder Elektronenzustand der Atome in ein Band auf, das maximal $2N$ Elektronen aufnehmen kann. So entstehen aus einem s-Zustand oder s-Orbital (↑ 4.2.3.2.) ein s-Band, aus den drei p-Zuständen einer Schale ein p-Band, das streng genommen drei Teilbänder mit je $2N$-Zuständen umfaßt, aus den d-Zuständen die d-Bänder usw. Für die Elektronen innerer Schalen ist die Überlappung der Eigenfunktionen jedoch gering; infolgedessen führt die Aufspaltung nur zu sehr schmalen Bändern, die normalerweise vollbesetzt sind. Bei den äußeren Bindungselektronen ergeben sich für die Breite eines Bands dagegen Werte in der Größenordnung von 1 eV, und zwischen den aus s-, p- oder d-Atomzuständen der gleichen Schale resultierenden Bändern tritt häufig Überlappung ein. Der Auffüllungsgrad des obersten, noch mit Elektronen besetzten Bands hängt von der Anzahl der Bindungselektronen sowie von einer eventuellen *Überlappung der Bänder* ab. Im Fall der Überlappung treten Elektronen aus besetzten in unbesetzte oder teilweise gefüllte Bänder über, bis die Fermienergie in allen Teilbändern den gleichen Wert erreicht.

8.2.1.2.3. *Blochfunktionen und Berechnung der Bandstruktur.* Nach einem grundlegenden Theorem, das von Bloch aufgestellt wurde, lassen sich die Lösungen der Schrödingergleichung im periodischen Potential eines Kristalls als ebene Wellen darstellen, deren Amplituden in der Umgebung der Gitterpunkte, also der Atomrümpfe, gleichartig moduliert sind. In einer Dimension besitzt eine solche Blochfunktion die Form

(6) $\psi_k(x) = u_k(x) \exp(ikx);$
$u_k(x) = u_k(x + na).$

Hier beschreibt das Exponentialglied eine in x-Richtung fortlaufende Welle, wobei der zeitliche Anteil $\exp(i\omega t)$ fortgelassen wurde (s. a. 8.2.1.1.). Die Amplitudenfunktion $u_k(x)$ soll gitterperiodisch verlaufen, d. h., sie wiederholt sich in Abständen vom ganzzahligen Vielfachen einer Gitterkonstanten a. Den Verlauf einer derartigen Welle und ihrer Anteile veranschaulicht Abb. 8.2.1-8. Physikalisch besagt das *Blochsche Theorem*, daß die mit der Wellenzahl k, bzw. in drei Dimensionen dem Wellenzahlvektor $\boldsymbol{k}$, indizierten Eigenfunktionen in den Gebieten zwischen den Atomrümpfen noch eine gewisse Ähnlichkeit zum Modell freier Elektronen im Potentialkasten aufweisen, während im Einflußbereich der Rumpfpotentiale starke Abweichungen zu erwarten sind.

Zur Berechnung der Bandstruktur in Gestalt der $E(\boldsymbol{k})$-Funktionen für die einzelnen Bänder wurden verschiedenartige Näherungsverfahren ausgearbeitet. Bewährt hat sich vor allem die Einführung von *Pseudo-* oder *Modellpotentialen*, in denen die Beeinflussung der Elektronen durch die Rumpfpotentiale sowie durch die übrigen Kristallelektronen zusammengefaßt wird. Als Lösungsansätze für die Schrödingergleichung benutzt man häufig sogenannte orthogonalisierte Planwellen (auch als *OPW-Methode* bezeichnet), die aus Anteilen ebener Wellen und den Eigenfunktionen der Atome aufgebaut werden und den Periodizitätsforderungen der Blochwellen [8.2.(7)] entsprechen. Unter Einsatz der elektronischen Rechentechnik ist es mit zunehmender Genauigkeit gelungen, die Bandstruktur konkreter Festkörper (Halbleiter, Metalle) quantiativ zu berechnen.

8.2.1.3. Darstellungsformen des Bändermodells

8.2.1.3.1. *Bändermodell im Ortsraum.* Bisher wurden die Energiezustände der Kristallelektronen im $\boldsymbol{k}$-Raum betrachtet. Diese Darstellungsweise bietet den Vorteil, daß der Zusammenhang zwischen der Elektronenenergie und

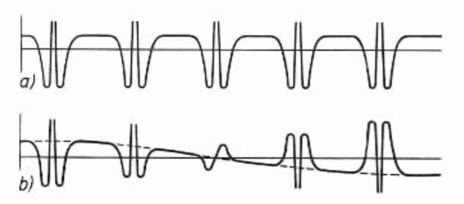
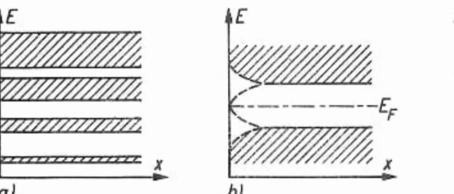


Abb. 8.2.1-8 Blochsche Elektronenwelle im Eindimensionalen; a) gitterperiodische Amplitudenfunktion $u_k(x)$, b) Blochsche Welle $\psi_k(x)$, ebene Welle *gestrichelt*

dem Wellenzahlvektor $\boldsymbol{k}$, der als Maß für den Elektronenimpuls aufgefaßt werden kann, unmittelbar sichtbar wird. Für viele Zwecke reicht es jedoch aus, in den gewohnten Ortsraum zurückzukehren und die Energiezustände als Funktion der Ortskoordinaten aufzutragen. In einer Dimension ergeben sich die erlaubten



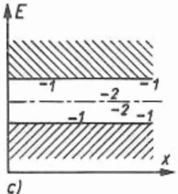


Abb. 8.2.1-9 a) Bändermodell im Ortsraum für einen eindimensionalen Kristall; *schraffiert* erlaubte Bereiche, *weiß* verbotene Zonen, *1* flache, *2* tieferliegende Störniveaus; b) Bandverbiegung an Oberflächen, c) Lage lokalisierter Elektronenzustände von Fremdatomen; das Ferminiveau E_F wird nicht abgebogen

Energiebereiche der Elektronen im idealen Kristall dann einfach als bandförmige Streifen parallel zur x-Achse (Abb. 8.2.1-9a). Bei dieser Darstellungsweise, die den Namen *Bändermodell* plastisch zum Ausdruck bringt, ist der Zusammenhang zwischen der Energie und dem Impuls der Kristallelektronen nicht zu erkennen; man sollte daher stets daran denken, daß zu jedem der diskreten Zustände eines Bands ganz bestimmte Werte des Wellenzahlvektors bzw. des Impulses gehören. Andererseits lassen sich örtliche Veränderungen der Elektronenzustände leicht berücksichtigen. So treten Sprünge in der Banddarstellung auf, wenn verschiedene Materialien aneinander grenzen, und an Festkörperoberflächen kommt es durch den Einfluß elektrischer Felder häufig zu *Bandverbiegungen* (Abb. 8.2.1-9b). Wichtig ist ferner das Auftreten von *lokalisierten Elektronenzuständen* innerhalb der verbotenen Zonen bei Abweichungen vom idealen Kristallbau. Beispielsweise bewirkt der Einbau von Fremdatomen in den Festkörpern meist das Zustandekommen von diskreten Zuständen, deren Energieniveaus mehr oder weniger weit von den Bandkanten entfernt im verbotenen Bereich liegen. Die entsprechenden Eigenfunktionen sind im Unterschied zu den Bandzuständen nur in der Nähe der Fremdatome von Null verschieden. In der Ortsdarstellung des Bändermodells werden solche Zustände durch Striche veranschaulicht (Abb. 8.2.1-9c).

8.2.1.3.2. *Bändermodell im reduzierten Zonenschema.* Bei der Darstellung im $\boldsymbol{k}$-Raum tritt die Frage auf, ob die getroffene Zuordnung der Bandzustände zu den verschiedenen Brillouinzonen zwingend ist. Wegen der Gitterperiodizität der Blochfunktionen [8.2.(6)] ist das nicht der Fall, denn der Verlauf der Elektronenwellen ändert sich nicht, wenn zur eindimensionalen Wellenzahl k ganze Vielfache von $2\pi/a$ hinzugefügt werden. In drei Dimensionen ist entsprechend eine Addition oder Subtraktion eines beliebigen Gittervektors $\boldsymbol{G}$ des reziproken Gitters zum bzw. vom Wellenzahlvektor $\boldsymbol{k}$ ohne Änderung des physikalischen Gehalts erlaubt. Infolgedessen ist es zulässig, die Elektronenzustände der höheren Bänder sämtlich auf die 1. Brillouinzone abzubilden oder, wie man sagt, zu reduzieren. Im eindimensionalen Fall ist dazu dasjenige Vielfache von $2\pi/a$ zu den Wellenzahlen hinzuzufügen oder davon abzuziehen, das den jeweiligen Ast der Funktion $E(k)$ in die 1. Brillouinzone transformiert (Abb. 8.2.1-10). Die so gewonnene Banddarstellung im reduzierten Zonenschema ist meist vorteilhaft, weil die Bänder dann zusammenhängende Gebilde werden. Aber auch unter physikalischen Gesichtspunkten lassen sich Argumente zugunsten der reduzierten Darstellung anführen. Die in der Näherung freier Elektronen bestehende Beziehung $\boldsymbol{p} = \hbar\boldsymbol{k}$ zwischen Wellenzahlvektor und Impuls gilt im Rahmen des Bändermodells nicht mehr. Vielmehr wird der Impuls an den Rändern der Brillouinzonen jeweils Null, da stehende Wellen ruhenden Elektronen entsprechen, und $\boldsymbol{k}$ stellt innerhalb der Bänder nur noch

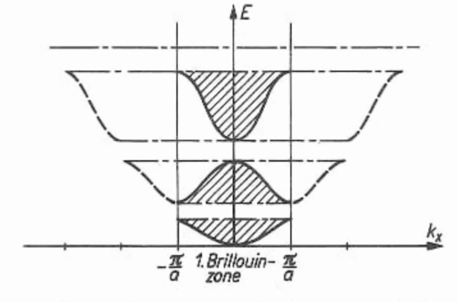


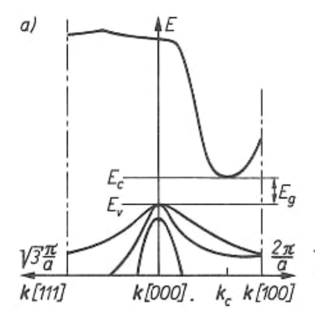
Abb. 8.2.1-10 Bändermodell im reduzierten Zonenschema

ein Maß für den *Pseudoimpuls des Kristallelektrons* dar. Ähnlich wie bei den Phononen (s. a. 8.1.2.2.) spielt dieser Pseudoimpuls eine wichtige Rolle als Erhaltungsgröße bei Wechselwirkungen zwischen Kristallelektronen und Phononen, Photonen oder anderen Quasiteilchen. Dabei sind stets nur Änderungen von $\boldsymbol{k}$, dagegen nicht die Absolutwerte, von Interesse.

8.2.1.3.3. *Bändermodell im periodischen Zonenschema.* Wenn es möglich ist, die Elektronenzustände der höheren Brillouinzonen in die 1. Zone zu reduzieren, so lassen sich natürlich auch umgekehrt die Zustände aus der 1. Brillouinzone auf die höheren Zonen abbilden. Man gelangt zu dieser Darstellung, indem der $E(\boldsymbol{k})$-Verlauf des reduzierten Zonenschemas für jedes Band in den höheren Brillouinzonen periodisch fortgeführt wird. Physikalisch bringt das so entstehende *periodische oder wiederholte Zonenschema* zum Ausdruck, daß die Energie der Kristallelektronen im reziproken Gitter einen periodischen Verlauf zeigt. Diese Darstellungsform erweist sich besonders bei der Untersuchung von Feinheiten der Bewegung von Metallelektronen in Magnetfeldern als vorteilhaft (↑ 8.4.4.2.).

für dreidimensionale Kristalle müßte demnach den Verlauf der erlaubten Energiewerte als Funktion des Wellenzahlvektors $E(\boldsymbol{k})$ wiedergeben. Eine anschauliche Darstellung dieser Funktion ist nicht ohne weiteres möglich, weil hierfür vier Dimensionen, nämlich E und die drei Komponenten von $\boldsymbol{k}$ in den Richtungen der Kristallachsen erforderlich wären. Man hilft sich in der Praxis, indem der $E(\boldsymbol{k})$-Verlauf für eindimensionale Schnitte längs wichtiger Richtungen im Kristall dargestellt wird. Abb. 8.2.1-11 veranschaulicht dies für die reale Bandstruktur der wichtigen Halbleitermaterialien Silicium bzw. Galliumarsenid. Man erkennt, daß der Bandverlauf in den ⟨100⟩-Richtungen parallel zu den Würfelkanten der kubischen Zellen (rechter Teil der Diagramme) wesentlich von dem Verlauf in den ⟨111⟩-Richtungen längs der Würfeldiagonalen abweicht.

Eine andere Möglichkeit zur Veranschaulichung der Elektronenzustände in dreidimensionalen Kristallen besteht darin, daß man nicht die $E(\boldsymbol{k})$-Funktionen selbst, sondern nur die Flächen konstanter Elektronenenergie $E(\boldsymbol{k}) = \text{const}$ darstellt. Für das Modell freier Elektronen besitzen diese Flächen die Gestalt konzentrischer



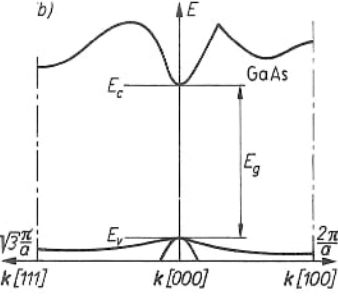


Abb. 8.2.1-11 Bandstruktur wichtiger Halbleiter für verschiedene Richtungen im Kristallgitter; a) Silicium, b) Galliumarsenid; E_c, E_v Bandkanten

8.2.1.4. Reale Bandstrukturen

8.2.1.4.1. *Bändermodell in drei Dimensionen.* In den vorausgegangenen Abschnitten wurden die prinzipiellen Aussagen des Bändermodells der Elektronen im Festkörper der Einfachheit halber vorwiegend am eindimensionalen Modell einer linearen Atomkette begründet und diskutiert. Bei der Erweiterung zu den drei Dimensionen eines räumlich ausgedehnten Kristalls muß beachtet werden, daß der Potentialverlauf zwischen den Atomrümpfen wegen der gerichteten Bindungsanteile und der unterschiedlichen Atomabstände in den verschiedenen Netzebenen von der betrachteten Richtung abhängt, also anisotrop ist. Da der Potentialverlauf den Verlauf der $E(\boldsymbol{k})$-Funktion in den einzelnen Bändern bestimmt, resultiert folglich eine Bandstruktur, die für unterschiedliche Kristallrichtungen wesentlich verschieden sein kann. Ein vollständiges Elektronenbandschema

Kugeln, während für wirkliche Kristalle Flächen komplizierter Gestalt erhalten werden (s. a. 8.4.4.2.). Vielfach sind diese Flächen nicht zusammenhängend, wie Abb. 8.2.1-11 für Energieflächen im unteren Teil des Leitungsbands (↑ 8.3.1.) von Silicium zeigt. Eine weitere Besonderheit des dreidimensionalen Bändermodells besteht in der Möglichkeit einer energetischen Überlappung der verschiedenen Bandzweige, die nach der Blochschen Näherung (↑ 8.2.1.2.2.) aus den Elektronenzuständen der Atome hervorgehen.

8.2.1.4.2. *Bandzweige.* Bei der Bandstruktur der obersten, mit Elektronen besetzten Bänder von Silicium, Galliumarsenid (↑ Abb. 8.2.1-11a, b) bezeichnet man das untere dieser beiden Bänder als *Valenzband*, das obere als *Leitungsband* (↑8.2.2.). Es fällt auf, daß die $E(\boldsymbol{k})$-Funktionen in mehrere Zweige gegliedert sind. Zu einem Energiezustand gehören dann mehrere mögliche

Beträge des Wellenzahlvektors oder verschiedene Bewegungsformen mit unterschiedlichem Impuls. Diese Aufspaltung der Bänder in Zweige kommt zustande, weil die Entartung der Atomzustände, beispielsweise der $6p$-Zustände (s. a. 4.2.4.), beim Übergang von den Elektronenzuständen einzelner Atome zu denen des Festkörpers teilweise aufgehoben wird. Im Falle des Leitungsbands berühren sich die Bandzweige auch für $\boldsymbol{k} = 0$ nicht alle, sondern es tritt eine kleine Energiedifferenz auf, die eine Folge der *Spin-Bahn-Kopplung*, d. h. der wechselseitigen Beeinflussung zwischen Spin- und Bahnbewegung der Elektronen, darstellt.

8.2.1.4.3. *Extremwerte der Elektronenenergie.* Bei genügend starkem Einfluß des periodischen Gitterpotentials ergeben sich schließlich immer stärkere Abweichungen der $E(\boldsymbol{k})$-Funktionen vom parabolischen Verlauf des Modells freier Elektronen (↑ Abb. 8.2.1-6). Bei den Bandstrukturen von Si, GaAs wird dies vor allem für das Leitungsband deutlich (↑ Abb. 8.2.1-11 a, b). So liegt das Minimum der Energie oder die untere Bandkante im Falle von Silicium nicht bei $\boldsymbol{k} = 0$, sondern in den $\langle 100 \rangle$-Richtungen bei $0{,}85\, k_1$, wenn k_1 den Rand der ersten Brillouinzone kennzeichnet. Das Maximum der Elektronenenergie im Valenzband (bei $\boldsymbol{k} = 0$) und das Minimum im Leitungsband entsprechen dann verschiedenen $\boldsymbol{k}$-Werten, und ein Übergang von einem Elektron aus dem oberen Bereich eines Bands in den unteren des darüberliegenden ist nur bei gleichzeitiger, beträchtlicher Änderung des Impulses möglich. Dieser Sachverhalt spielt bei *optischen Elektronenübergängen* in Halbleitern eine wichtige Rolle (s. a. 8.8.2.). Halbleiter vom Typ des Siliciums, bei denen der minimale Bandabstand E_g verschiedenen $\boldsymbol{k}$-Werten entspricht, werden als *indirekte Halbleiter* bezeichnet. Im Unterschied dazu weist die $E(\boldsymbol{k})$-Funktion für GaAs zwar Minima bei verschiedenen $\boldsymbol{k}$-Werten auf (Vieltalstruktur), jedoch liegt das tiefere Minimum bei $\boldsymbol{k} = 0$. Infolgedessen entspricht der Bandabstand E_g Übergängen zum gleichen $\boldsymbol{k}$-Wert, und GaAs ist ein Vertreter der *direkten Halbleiter*. Die hier diskutierten Feinheiten des Bändermodells sind von großer Bedeutung für die genaue Beschreibung des elektronischen und optischen Verhaltens von Festkörpern. Eine näherungsweise Behandlung der elektrischen Leitungsprozesse, insbesondere auch der wichtigsten Vorgänge in Halbleitern, ist jedoch schon auf der Basis des eindimensionalen Bändermodells unter Übernahme von Beziehungen des Modells freier Elektronen möglich, wovon im folgenden Gebrauch gemacht wird. Ein derartiges Vorgehen ist gerechtfertigt, weil die $E(\boldsymbol{k})$-Bereiche in der Umgebung von Maxima bzw. Minima der Energie durch Paraboloide angenähert werden können, die sich der realen Verteilung der Energiezustände örtlich anschmiegen.

8.2.2. Besetzung der Elektronenzustände

8.2.2.1. Besetzung bei $T = 0$ K. Ausgehend von der Beschreibung möglicher Elektronenzustände im Festkörper ist zu klären, in welcher Weise diese Zustände tatsächlich mit Elektronen besetzt sind. Betrachtet sei zunächst ein Festkörper am absoluten Nullpunkt $T = 0$ K, dem auch keine Energie in Form von Teilchen- oder Wellenstrahlen zugeführt wird. Nach den Gesetzmäßigkeiten der klassischen Physik sollten die Kristallelektronen unter diesen Bedingungen kondensieren, d. h. im tiefsten Energiezustand zur Ruhe kommen. Für Elektronen als Teilchen mit halbzahligem Spin oder Fermionen gilt jedoch das *Pauliprinzip* der Quantentheorie (↑ 2.5.5.2.). Danach können sich höchstens zwei Elektronen, deren Spins entgegengesetzt orientiert sind, in einem durch die Komponenten des Wellenzahlvektors charakterisierten Zustand aufhalten. Falls das Band mehrere Zweige umfaßt, multipliziert sich diese Zahl entsprechend, weil die zu verschiedenen Zweigen gehörenden Zustände gleicher Energie getrennt zu rechnen sind. Auf jeden Fall ordnen sich die Elektronen im Festkörper am absoluten Nullpunkt so an, daß sämtliche Zustände unterhalb einer Grenzenergie $E_F(0)$ vollständig aufgefüllt sind, während die darüberliegenden Zustände leer bleiben. Die Grenzenergie $E_F(0)$ wird *Fermienergie* bei $T = 0$ K genannt. Im

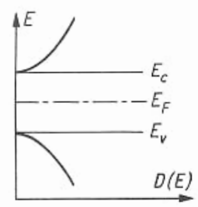


Abb. 8.2.2-1 Parabolische Näherung für die Zustandsdichtefunktion $D(E)$ nahe der Bandkanten; gegenüber Abb. 8.2.1-4 wurden die Achsen vertauscht, um die Lage im Bändermodell zu veranschaulichen; E_F Fermienergie, E_c, E_v Bandkanten

Rahmen der Näherung freier Elektronen wurde dieses besondere Verhalten der Elektronen im Festkörper bereits früher betrachtet (↑ 8.2.1.1.). Die dort begründete parabolische Beziehung [8.2.(5)] für die *Zustandsdichte* $D(E)$ läßt sich sogar in das Bändermodell übernehmen, sofern nur Zustände in der Nähe der Bandkanten betrachtet werden. Man setzt dann in brauchbarer

Näherung

(7) $D(E) = 1/2\pi^2(2m^*/\hbar^2)^{3/2}(E - E_c)^2$;
$D(E) = 1/2\pi^2(2m^*/\hbar^2)^{3/2}(E_v - E)^2$

für die Zustandsdichte nahe der unteren Kante des Leitungsbands E_c bzw. nahe der oberen Kante des Valenzbands E_v (Abb. 8.2.2-1). In diesen Formeln tritt somit die Energiedifferenz zu den Kanten E_c bzw. E_v der Bänder auf, die der Fermienergie benachbart sind. Ferner werden Abweichungen von der Näherung freier Elektronen dadurch berücksichtigt, daß die Masse m_e des freien Elektrons durch eine *effektive Masse* m^* ersetzt wird, die im allgemeinen selbst von der Energie abhängt und empirisch bestimmt werden kann. Im Inneren der Bänder

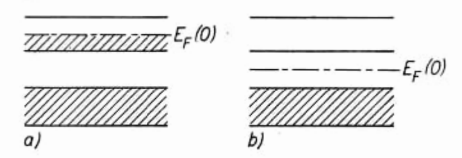


Abb. 8.2.2-2 Lage des Ferminiveaus $E_F(0)$ und Bandbesetzung bei $T = 0$ K; a) für Metalle, b) für Halbleiter und Isolatoren

besitzt die Zustandsdichte freilich einen komplizierteren Verlauf. Hinsichtlich der Besetzung der Elektronenzustände bei $T = 0$ K sind zwei wichtige Fälle zu unterscheiden. Entweder liegt $E_F(0)$ innerhalb eines Bands erlaubter Elektronenzustände oder im verbotenen Energiebereich zwischen dem Valenz- und dem Leitungsband (Abb. 8.2.2-2). Im ersten Fall ist das oberste Band, das noch Elektronen enthält, nur teilweise gefüllt, während im zweiten Falle bei $T = 0$ K ein leeres Band über einem vollbesetzten Band liegt. Wie im folgenden noch begründet wird, entspricht ein teilweise besetztes oberes Band den Gegebenheiten metallischer Leiter, während bei Halbleitern und Isolatoren das oberste noch Elektronen enthaltende Band am absoluten Nullpunkt vollständig gefüllt ist.

Für das Verständnis der Elektronen im Festkörper ist es wichtig, daß die besetzten Elektronenzustände unterhalb $E_F(0)$ nicht etwa ruhenden, sondern im Sinne der Blochfunktionen [8.2.(6)] bewegten Elektronen entsprechen. Selbst bei $T = 0$ K befinden sich die Elektronen im Festkörper somit in lebhafter Bewegung; ihre kinetische Energie kann mit Werten bis zu einigen Elektronenvolt sogar Beträge erreichen, die weit über der Energie der Wärmebewegung von Atomen bei Temperaturen unterhalb des Schmelzpunkts fester Stoffe liegt.

8.2.2.2. Statistik der Kristallelektronen. Wird ein Festkörper beginnend von $T = 0$ K allmählich aufgeheizt, so sollte man erwarten, daß die Kristallelektronen an der zunehmend angefachten Wärmebewegung teilnehmen, indem sie durch *Elektron-Phonon-Wechselwirkung* (s. a. 8.1.2.5.) und durch *Elektron-Elektron-Wechselwirkung* sowohl mit den Gitterbausteinen als auch untereinander Energie austauschen. Tatsächlich nimmt jedoch nur ein kleiner Teil der Kristallelektronen beim Aufheizen zusätzliche Energie auf, wie beispielsweise aus dem geringen Beitrag der Elektronen zur *spezifischen Wärme der Festkörper* hervorgeht (↑ 8.4.3.5.). Dies ist dadurch bedingt, daß die Breite des Energiebereichs der besetzten Bandzustände einige Elektronenvolt beträgt, während die Wärmeenergie pro Teilchen bei der Temperatur T nur von der Größenordnung k_BT ist und bei 300 K erst rund 0,025 eV erreicht. Infolgedessen können Kristallelektronen, die sich in den Bandzuständen wesentlich unterhalb der Fermigrenze aufhalten, bei Wechselwirkungsprozessen gar keine Energie aufnehmen oder abgeben, weil die Zustände, in die sie dabei übergehen müßten, alle schon besetzt sind. Eine Ausnahme bilden die Elektronen, die sich in einem relativ schmalen Energiebereich von der Größenordnung k_BT unterhalb der Fermienergie aufhalten, denn diese Elektronen können durch Aufnahme von Wärmeenergie in die unbesetzten Zustände oberhalb $E_F(0)$ gelangen. Ist das betrachtete Band nur teilweise gefüllt, so wird die ursprünglich scharfe Fermigrenze zwischen besetzten und unbesetzten Zuständen verwischt. Falls das oberste besetzte Band vollständig gefüllt ist, sind die nächsten unbesetzten Zustände – abgesehen von lokalisierten Niveaus in der verbotenen Zone – freilich erst im darüberliegenden leeren Band verfügbar; eine thermische Anregung von Elektronen vom Valenz- in das Leitungsband ist deshalb nur möglich, wenn der Bandabstand E_g im Vergleich zur mittleren Wärmeenergie k_BT nicht zu groß ist.

Zusammenfassend kann festgestellt werden, daß sich das Elektronengas im Festkörper aufgrund der Gesetzmäßigkeiten der Quantentheorie in seinem thermischen Verhalten sehr stark von dem eines klassischen Teilchengases unterscheidet; man bezeichnet es daher als *entartet*.

Quantitativ wird die Wahrscheinlichkeit für die Besetzung der Elektronenzustände im thermischen Gleichgewicht durch die *Fermi-Dirac-Statistik* beschrieben. Bei ihrer Begründung geht man, ebenso wie bei der im Zusammenhang mit den Phononen behandelten *Bose-Einstein-Statistik* (↑ 8.1.3.2.), vom Prinzip der Nichtunterscheidbarkeit atomarer Teilchen aus, d. h., Vertauschungen von Elektronen untereinander werden nicht als verschiedene Mikrozustände des Systems betrachtet. Zusätzlich ist jedoch dem Pauliprinzip Rechnung zu tragen, nach dem jeder durch die drei Komponenten des Wellenzahlvektors charakterisierte Zustand höchstens zwei Elektronen mit entgegengerichte-

tem Spin aufnehmen darf. Die Rechnung führt auf die *Verteilungsfunktion* von FERMI und DIRAC:

(8) $f = 1/[\exp(E - E_F)/k_B T + 1]$.

Physikalisch gibt diese wichtige Funktion die Wahrscheinlichkeit dafür an, daß ein bei der Energie E liegender Zustand mit einem Elektron besetzt ist; entsprechend ist $(1 - f)$ die Wahrscheinlichkeit dafür, daß der betrachtete

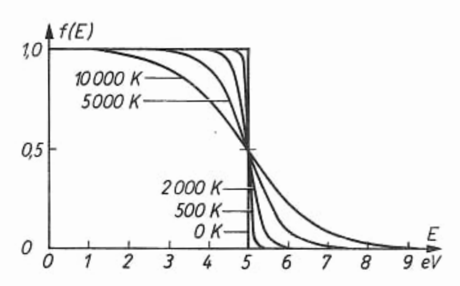


Abb. 8.2.2-3 Fermi-Dirac-Verteilungsfunktion für verschiedene Temperaturen mit $E_F \approx E_F(0) = 5$ eV

Zustand unbesetzt ist. Der Verlauf dieser Verteilungsfunktion ist in Abb. 8.2.2-3 für verschiedene Werte des Verhältnisses $(E_F/k_B T)$ aufgetragen. Die hier als Materialgröße auftretende Fermienergie E_F hängt selbst geringfügig von der Temperatur T ab und ist etwas kleiner als die scharfe Grenze $E_F(0)$ zwischen besetzten und unbesetzten Zuständen bei $T = 0$ K. Setzt man in [8.2.(8)] $E = E_F$, so wird $f = 0{,}5$. Physikalisch entspricht die *Fermienergie bei beliebigem T*, die man auch als *chemisches Potential* bezeichnet, somit derjenigen Energie, für die die Wahrscheinlichkeit, ein Elektron im betrachteten Zustand anzutreffen, ebenso groß ist wie die Wahrscheinlichkeit, daß der Zustand unbesetzt ist. Dabei wird zunächst vorausgesetzt, daß jeder Zustand nur mit einem Elektron besetzt sein kann; die doppelte Besetzungsmöglichkeit durch Elektronen entgegengesetzten Spins wird bei der Abzählung der Zustände mittels der Zustandsdichtefunktion [8.2.(7)] durch einen darin schon enthaltenen Gewichtsfaktor 2 berücksichtigt.

Für den Grenzfall $T = 0$ K ergibt sich für die Verteilungsfunktion $f(E)$ ein Rechteckverlauf. Mit wachsender Temperatur wird die Breite des Gebiets, in dem f von den Werten 1 bzw. 0 abweicht, immer größer. Trotzdem bleibt dieser Bereich, gemessen an der gesamten Breite der Energieverteilung der Kristallelektronen für feste Stoffe, auch bei hohen Temperaturen relativ begrenzt, denn E_F beträgt mehrere Elektronenvolt, so daß $E_F/k_B T$ stets viel kleiner als 1 bleibt. Theoretisch würde die Fermi-Dirac-Verteilung bei sehr hohen Temperaturen in die klassische Verteilungsfunktion der *Boltzmannstatistik* übergehen (↑ 2.4.4. und 6.1.3.), jedoch ist dies praktisch bedeutungslos, weil bei Temperaturen in der Größenordnung von 10^5 bis 10^6 K kein Stoff mehr im festen Zustand vorliegt. Für viele Rechnungen ist es jedoch nützlich, daß die 1 im Nenner von [8.2.(8)] vernachlässigt werden kann, wenn die Differenz $(E - E_F)$ größer als einige $k_B T$ ist, also nur energiereiche Elektronen betrachtet werden. Für den Ausläufer des Spektrums der Kristallelektronen bei relativ hohen Anregungsenergien gilt dann in guter Näherung ein zur Boltzmannstatistik analoger Exponentialverlauf, d. h., die Entartung des Elektronengases ist in diesem Bereich weitgehend aufgehoben.

8.2.2.3. Berechnung der Elektronenbesetzung. Die Anzahl der Elektronen pro Volumeneinheit dN, die sich im Festkörper bei gegebener Temperatur T in dem differentiellen Energiebereich zwischen E und $E + dE$ aufhalten, ergibt sich durch Multiplikation der Fermi-Dirac-Verteilungsfunktion $f(E)$ mit der Zustandsdichtefunktion $D(E)$:

(9) $dN = D(E) f(E)\, dE$.

Um die Zahl der Elektronen in einem Energiebereich endlicher Breite, beispielsweise in einem ganzen Band, für eine bestimmte Temperatur bei thermischem Gleichgewicht zu berechnen, ist das Integral von [8.2.(9)] über das betrachtete Energiegebiet zu bilden. Bei der Anwendung auf einen konkreten Festkörper muß man dazu die Fermienergie E_F sowie die Zustandsdichtefunktion $D(E)$ kennen. E_F läßt sich beispielsweise durch Messung der *Elektronenaustrittsarbeit* (↑ 8.4.3.6.) oder aus Leitfähigkeitsuntersuchungen ermitteln, für $D(E)$ benutzt man theoretisch berechnete Ausdrücke oder im Falle von Halbleitern die parabolischen Näherungen entsprechend [8.2.(7)]. Die in den letztgenannten Ausdrücken enthaltenen effektiven Elektronenmassen sowie die Lage der Bandkanten, die man auch für die Festlegung der Integrationsgrenzen benötigt, gewinnt man aus elektrischen und optischen Meßdaten.

Die hier dargelegten Gesetzmäßigkeiten lassen sich im Rahmen eines erweiterten Bändermodells auch auf reale Festkörper mit Kristallbaufehlern und Verunreinigungen anwenden. In der verbotenen Zone zwischen Valenz- und Leitungsband treten dann lokalisierte Elektronenniveaus zusätzlich zu den Bandzuständen auf, was durch spezielle Beiträge zur Zustandsdichtefunktion berücksichtigt werden kann. Der Einfluß solcher Störterme auf die Besetzung der Bänder mit Elektronen ist für das Verständnis der elektrischen Leitungsvorgänge in Halbleitern und der optischen Eigenschaften

von Festkörpern von grundlegender Bedeutung (s. a. 8.3. und 8.8.).

8.2.3. Dynamik der Kristallelektronen

8.2.3.1. Leiter, Isolatoren, Halbleiter. Die prinzipielle Gültigkeit des Bändermodells der Kristallelektronen zeigt sich deutlich, wenn man die Bewegung der Elektronen im Festkörper unter dem Einfluß innerer oder äußerer Ursachen untersucht. Beispiele dynamischer Elektronenphänomene sind die elektrische Leitfähigkeit, der Elektronenbeitrag zur thermischen Leitfähigkeit, die Bewegung der Kristallelektronen in magnetischen Feldern oder ihre Wechselwirkung mit Photonen bzw. Phononen. Betrachtet sei zunächst die besonders wichtige *elektrische Leitfähigkeit*. Wenn an einen Festkörper eine elektrische Spannung angelegt wird, so baut sich im Inneren eine elektrische Feldstärke $\boldsymbol{E}$ auf, und auf ein Kristallelektron wirkt die zusätzliche Kraft $-e\boldsymbol{E}$ ein (e positive Elementarladung). In der Näherung freier Elektronen müßten die Elektronen beschleunigt werden, also zusätzliche kinetische Energie erlangen, wodurch sie in höhere Energiezustände gelangen. Die zusätzliche Elektronenbewegung folgt natürlich der Feldrichtung; die Elektronen bewegen sich dann mit einer Driftgeschwindigkeit v_D, die sich der ohne Vorzugsrichtung erfolgenden Bewegung ohne Feld überlagert. Hierdurch kommt es zu einem Elektronen- oder Ladungstransport, d. h., es fließt ein elektrischer Strom. Für das teilweise mit Elektronen gefüllte Valenzband eines metallischen Leiters bleibt diese Überlegung zumindest qualitativ gültig. Abb. 8.2.3-1 deutet an, wie sich die gesamte Fermikugel der besetzten Elektronenzustände im $\boldsymbol{k}$-Raum bei Anlegen eines elektrischen Felds um ein Stück verschiebt. Im Endeffekt wird dadurch, wie die schraffierten Bereiche ausweisen, die Verteilung der Elektronen

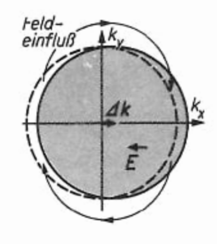


Abb. 8.2.3-1 Verschiebung der Fermikugel um $\Delta\boldsymbol{k}$ bei Anlegen eines elektrischen Felds E

über die Energiezustände nur in den sichelförmigen Bereichen nahe der Fermienergie verändert. Für diese Elektronen gibt es zu einem Wellenzahlvektor $\boldsymbol{k}$ kein Gegenstück mit dem Vektor $-\boldsymbol{k}$, so daß ein Nettostromfluß resultiert. Aus Abschätzungen folgt, daß die Energieaufnahme der Elektronen aus dem Feld unter üblichen Bedingungen sehr klein im Vergleich zur mittleren Elektronenenergie ohne Feld ist; demnach wird die Verteilung der besetzten Elektronenzustände nur geringfügig verändert. Daraus ergeben sich schwerwiegende Konsequenzen für das Verhalten der Elektronen in vollbesetzten Energiebändern. Da die Energiezufuhr durch das elektrische Feld nicht ausreicht, um Elektronen in ein höheres unbesetztes Band unter Überspringen der verbotenen Zone anzuregen, innerhalb des Bands aber keine freien Zustände verfügbar sind, kann sich die Bewegung der Gesamtheit der Elektronen unter dem Einfluß des Felds überhaupt nicht verändern. Andererseits folgt aus sehr allgemeinen Erwägungen zur Gittersymmetrie, daß die Bandzustände bezüglich einer Inversion von $\boldsymbol{k}$ symmetrisch verteilt sind. Zu jedem Zustand $\boldsymbol{k}$ gibt es einen äquivalenten Zustand $-\boldsymbol{k}$, und bei vollständiger Besetzung müssen sich die Anteile der Kristallelektronen zum Ladungstransport kompensieren. Für ein vollbesetztes Band gilt allgemein $\sum \boldsymbol{k} = 0$. Demzufolge kann ein vollbesetztes Band nichts zum Stromfluß beitragen!

Festkörper, bei denen die Bindungselektronen das Valenzband vollständig auffüllen, können am absoluten Nullpunkt der Temperatur daher den elektrischen Strom nicht leiten. Das Verhalten bei höheren Temperaturen wird durch die Größe des Bandabstands E_g zum nächsthöheren leeren Band bestimmt. Falls E_g nicht zu groß ist, können Elektronen gemäß der Fermi-Dirac-Statistik [8.2.(8)] durch thermische Anregung mit einer gewissen Wahrscheinlichkeit in das höhere Band gelangen. Die Elektronen in diesem Leitungsband, aber auch die im Valenzband zurückbleibenden unbesetzten Zustände oder Löcher ermöglichen dann eine Stromleitung. Da die Leitfähigkeit gegenüber metallischen Leitern jedoch deutlich geringer ist und bei $T = 0$ K sogar nahezu verschwindet, werden solche Festkörper mit einem Bandabstand bis zu etwa 2,5 eV als Halbleiter bezeichnet. Die Isolatoren unterscheiden sich von den Halbleitern nicht prinzipiell, sondern nur graduell durch ihren höheren Bandabstand E_g, der Werte von 2,5 bis 12 eV annehmen kann. Da E_g exponentiell in die Fermi-Dirac-Verteilungsfunktion eingeht, ist dann bei Zimmertemperatur nur noch eine sehr geringe Leitfähigkeit möglich. Im übrigen wird die Unterscheidung zwischen Halbleiter und Isolator nicht immer streng vorgenommen. Beispielsweise ordnet man ZnO mit $E_g = 3{,}2$ eV in Tabellen einmal als Isolator, ein anderes Mal als Halbleiter ein, je nachdem ob die noch gut meßbaren Leitungseigenschaften oder das dielektrische Verhalten betrachtet werden. Der Unterschied zwischen metallischen Leitern und Nicht- bzw. Halblei-

tern ist dagegen im Prinzip klar definiert durch das Nichtvorhanden- oder Vorhandensein einer endlichen Bandlücke E_g bei $T = 0$ K. Freilich gibt es auch hier Grenzfälle, von denen in jüngster Zeit der *interkalierte Graphit* besonderes Interesse erweckt hat. Die Kristalle des Graphits bestehen aus Schichten mit hexagonaler Anordnung der Kohlenstoffatome, die Bindung zwischen den Schichten erfolgt nur über relativ schwache van-der-Waals-Kräfte (s. a. 8.4.1.4.). Infolge dieser Struktur weist Graphit eine stark anisotrope Leitfähigkeit auf. Parallel zu den Schichtebenen kann die spezifische Leitfähigkeit bei $T = 300$ K etwa 1/25 des Werts von Kupfer erreichen, während sie senkrecht zu den Schichtebenen, in Richtung der c-Achse, um Größenordnungen geringer ist. Durch geeignete Behandlung in Lösungen oder Schmelzen läßt sich das Graphitgitter aufweiten, wobei zwischen C-Atom-Ebenen in regelmäßiger Folge nach jeder n-ten Schicht eine Lage anderer Teilchen, beispielsweise Li, K oder AsF_5, eingeschoben wird. Längs der Schichtebenen wurden bei diesem interkalierten Graphit Leitfähigkeiten gemessen, die den Werten der besten metallischen Leiter gleichkommen. Dieses Beispiel aktueller Forschung zeigt, daß die früher starr erscheinenden Grenzen zwischen Leitern und Halbleitern durch die Erforschung von Grenzfällen mit möglicherweise erheblicher künftiger Bedeutung aufgelockert werden.

8.2.3.2. Bewegungsgleichung der Kristallelektronen. Die quantitative Berechnung der Dynamik von Kristallelektronen bereitet Schwierigkeiten, weil äußere Felder nicht auf das Elektron allein einwirken, sondern auf das Gesamtsystem bestehend aus den Kristallelektronen und dem Gitter der Atomrümpfe. Dabei kann das betrachtete Elektron Energie oder Impuls an das Gitter übertragen oder umgekehrt vom Gitter aufnehmen. Diese Wechselwirkung modifiziert die Wirkung der äußeren Kräfte unter Umständen erheblich und kann sie hinsichtlich des Vorzeichens sogar umkehren. Glücklicherweise ist es bei Kenntnis des $E(\boldsymbol{k})$-Verlaufs im betrachteten Band möglich, die Bewegung der Kristallelektronen zu beschreiben, ohne die Einzelheiten der Wechselwirkung zwischen Elektronen und Gitter zu kennen.

Um eine Bewegungsgleichung für Kristallelektronen aufzustellen, bauen wir gedanklich ein lokalisiertes Elektron aus einem Wellenpaket von Eigenfunktionen des $\boldsymbol{k}$-Bereichs von $\boldsymbol{k}$ bis $\boldsymbol{k} + \Delta\boldsymbol{k}$ auf. Die Geschwindigkeit des Kristallelektrons entspricht bei dieser quantenmechanischen Beschreibung der *Gruppengeschwindigkeit* $v_k = \mathrm{d}\omega/\mathrm{d}k$ des Wellenpakets (↑ 1.2.3.). Wegen $E = \hbar\omega$ gilt $v_k = \hbar^{-1}\,\mathrm{d}E/\mathrm{d}k$. Nun bedeutet der Differentialquotient der Funktion $E(\boldsymbol{k})$ geometrisch bekanntlich die Steigung. Demnach läßt sich der relative Gang der Geschwindigkeit von

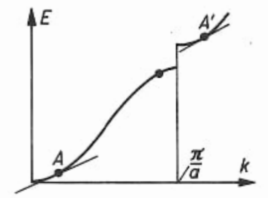
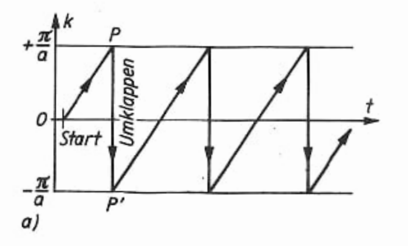


Abb. 8.2.3-2 Elektronengeschwindigkeit $\boldsymbol{v}$ in Bändern: die Gruppengeschwindigkeit eines Elektrons in den Zuständen bei A und A' entspricht der Steigung der an $E(\boldsymbol{k})$ angelegten Tangente, a Atomabstand

Kristallelektronen in den einzelnen Bandzuständen aus dem $E(\boldsymbol{k})$-Diagramm direkt ablesen (Abb. 8.2.3-2). Für den einfachsten Fall eines monotonen Verlaufs der Funktion $E(\boldsymbol{k})$ entnimmt man, daß v_k an der unteren Bandkante Null ist, danach bis zu einem Maximalwert ansteigt und schließlich bis zur oberen Bandkante wieder auf Null absinkt. Auflösen der obigen Beziehung für v_k liefert für die differentielle Änderung der Energie eines Kristallelektrons bei einer Änderung der Wellenzahl um dk die Beziehung $\mathrm{d}E = \hbar v_k\,\mathrm{d}k$. Durch Gleichsetzen mit der differentiellen Arbeit $-eEv_k\,\mathrm{d}t$ (entsprechend „Kraft $-eE$ mal Weg $v_k\,\mathrm{d}t$"), die in einem elektrischen Feld im Zeitintervall dt an das Elektron übertragen wird, erhält man als *Bewegungsgleichung des*



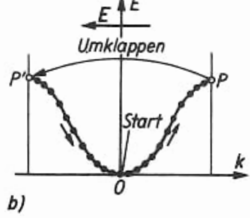


Abb. 8.2.3-3 Bewegung eines Kristallelektrons unter dem Einfluß einer konstanten äußeren Kraft (elektrisches Feld E); a) im $\boldsymbol{k}$-Raum (eindimensional), b) im $E(\boldsymbol{k})$-Diagramm

Kristallelektrons

(10) $\hbar \cdot \mathrm{d}\boldsymbol{k}/\mathrm{d}t = -e\boldsymbol{E}$.

Falls außer elektrischen auch magnetische Felder auf das Elektron einwirken, ist die rechte Seite von [8.2.(10)] durch Hinzufügen der *Lorentzkraft* $-e(\boldsymbol{v}_k \times \boldsymbol{B})$ zu erweitern (↑ 2.2.3.). Die so gewonnene Bewegungsgleichung beschreibt, wie sich der Pseudoimpuls $\hbar\boldsymbol{k}$ eines Kristallelektrons unter dem Einfluß äußerer Kräfte ändert. Wir wenden die Beziehung auf ein Elektron an, daß sich in einem sonst leeren Band eines Festkörpers befindet. Bei konstanter elektrischer Feldstärke $-\boldsymbol{E}$ soll auf das Elektron eine positive gleichbleibende Kraft $e\boldsymbol{E}$ einwirken. Wie Abb. 8.2.3-3 für den eindimensionalen Fall veranschaulicht, nimmt der Pseudoimpuls des betrachteten Elektrons linear zu, bis das Elektron im Zustand P an der oberen Bandkante anlangt. Dieser Zustand ist hinsichtlich des Pseudoimpulses jedoch physikalisch gleichwertig mit dem Zustand P' an der unteren Bandkante, da sich beide Positionen hinsichtlich $\boldsymbol{k}$ um einen reziproken Gittervektor, in einer Dimension um $2\pi/a$, unterscheiden (s. a. 8.2.1.3.2.). Infolgedessen kommt es zu einem „Umklappen" des Pseudoimpulses, wodurch das Elektron in die Position P' an der unteren Bandkante gelangt, von wo aus es bei konstanter Kraft erneut bis zur oberen Kante beschleunigt wird. Dieser Vorgang kann sich beliebig oft wiederholen. Ein Vergleich mit der $E(\boldsymbol{k})$-Kurve, die bei einem solchen Zyklus durchlaufen wird, führt zu einem merkwürdigen Ergebnis: Unter dem Einfluß einer äußeren Kraft wird ein Kristallelektron beschleunigt oder verzögert, je nachdem, ob es sich im unteren oder oberen Energiebereich des Bands aufhält. Dieser eigenartige Sachverhalt unterstreicht die eingangs getroffene Feststellung, wonach die Wechselwirkung des Kristallelektrons mit dem Gitter die Aktion der äußeren Kraft kompensieren oder sogar umkehren kann.

8.2.3.3. Effektive Masse der Kristallelektronen. Vielfach erweist es sich als zweckmäßig, die Dynamik von Kristallelektronen formal durch die klassische *Newtonsche Bewegungsgleichung* eines Teilchens (↑ 2.1.5.) zu beschreiben. Um die Wechselwirkung des Elektrons mit dem Gitter zu erfassen, muß man dann die träge Masse des Elektrons m_e durch eine *effektive Masse* m^* ersetzen, worauf früher bereits hingewiesen wurde (↑ 8.2.2.1.). In einer Raumdimension x schreiben wir als Bewegungsgleichung $(\mathrm{d}^2x/\mathrm{d}t^2) = 1/m^* eE_x$, d. h. „Beschleunigung gleich Kraft durch Masse". Der Zusammenhang zwischen der effektiven Masse m^* und der Funktion $E(\boldsymbol{k})$ für das betrachtete Band läßt sich leicht ableiten. Offensichtlich muß ja die Beschleunigung $(\mathrm{d}^2x/\mathrm{d}t^2)$ des Elektrons mit der Ableitung seiner Gruppengeschwindigkeit $\boldsymbol{v}_k = \hbar^{-1}\,\mathrm{d}E/\mathrm{d}\boldsymbol{k}$ nach der Zeit übereinstimmen, wenn die im vorigen Abschnitt eingeführte Bewegungsgleichung und die hier vorgenommene klassische Beschreibung im Einklang stehen sollen. Durch Differenzieren und Einsetzen ergibt sich $(\mathrm{d}^2x/\mathrm{d}t^2) = (\mathrm{d}\boldsymbol{v}_k/\mathrm{d}t) = \hbar^{-2}(\mathrm{d}^2E/\mathrm{d}\boldsymbol{k}^2)\,eE_x$ und aus dem Vergleich mit der oben angesetzten klassischen Bewegungsgleichung folgt für die effektive Masse

(11) $m^* = \hbar^2(\mathrm{d}^2E/\mathrm{d}\boldsymbol{k}^2)^{-1}$.

Auch diese wichtige Beziehung läßt sich an Hand des $E(\boldsymbol{k})$-Diagramms geometrisch interpretieren. Bekanntlich stellt die zweite Ableitung einer Funktion ein Maß für die Krümmung dar. Für das in Abb. 8.2.3-2 dargestellte Band ist die effektive Masse im unteren Teil positiv, weil die Krümmung hier gleichfalls positiv oder, von oben gesehen, konkav ist; im oberen Energiebereich des Bands treten wegen der konvexen Krümmung der $E(\boldsymbol{k})$-Kurve dagegen negative effektive Massen auf.

Zahlenwerte der effektiven Masse werden in Tabellen gewöhnlich durch das Verhältnis m^*/m_e ausgedrückt, das man auch Freiheitszahl des Kristallelektrons nennt. Für die Metalle Na, Cu, Ag, die sehr gute Elektrizitätsleiter sind, liegt das Verhältnis nahe 1, d. h., die Leitungselektronen in der Nähe der Fermienergie bewegen sich ähnlich wie freie Elektronen. Bei anderen Metallen und besonders bei Halbleitern zeigen sich viel größere Abweichungen, und es werden manchmal auch negative effektive Massen erhalten (Tab. 8.2.3-4). Für dreidimensionale Kristalle sind die Differentiationen in Formel 8.2.(11) nach den Raumkoordinaten auszuführen. Die effektive Masse wird dann mathematisch ein Tensor, also eine mathematische Größe, deren Betrag von der Bewegungsrichtung des Kristallelektrons abhängt. Betont sei nochmals, daß die tatsächliche träge Masse durch die Bindung der Elektronen an das Gitter natürlich nicht verändert wird; vielmehr bildet die Einführung der effektiven Masse nur ein

Tab. 8.2.3-4 Effektive Massen der Leitungselektronen für enige Metalle

Metall	Massenverhältnis m^*/m_e
Li	2,4
Na	1,3
Cu	1,5
Ag	1,0
Au	1,0
Be	0,46
Mg	1,43
Zn	0,9
Al	1,6
Pb	2,1

nützliches Hilfsmittel, um die Wechselwirkung der Kristallelektronen mit dem Gitter zu erfassen. Die so charakterisierten Elektronen bilden eine weitere Art von *Quasiteilchen*, denen physikalisch nur eine fiktive Bedeutung zukommt (s. a. 8.2.3.5.).

8.2.3.4. Messung der effektiven Masse. Um die effektive Masse experimentell zu bestimmen, geht man von elektrischen oder auch optischen Messungen aus. Allerdings können dabei auch für das gleiche Material unterschiedliche Werte gefunden werden, je nachdem welche Anteile des Kollektivs der Kristallelektronen vorzugsweise zum Meßeffekt beitragen. Man spricht in diesem Zusammenhang auch von elektrischen, thermischen oder optischen effektiven Massen.

Die Messung der elektrischen effektiven Masse erfolgt meist mit einem Effekt, der als *Zyklotronresonanz* bezeichnet wird, da hierbei eine gewisse Analogie zur Bewegung von Teilchen in einem Kreisbeschleuniger besteht (s. a. 8.4.4.2.). Die zu messende Probe wird, meist als dünne Schicht, in ein kräftiges Magnetfeld der konstanten Induktion $\boldsymbol{B}$ gebracht (Abb. 8.2.3-5). Senkrecht zu $\boldsymbol{B}$ wird durch Einstrahlung von Mikrowellen ein höchstfrequentes elektrisches Wechselfeld E erzeugt. Vom Magnetfeld werden die Kristallelektronen, die sich mit der Geschwindigkeit $\boldsymbol{v}_k$ bewegen, in Kreisbahnen abgelenkt, wobei für die Kreisfrequenz durch Gleichsetzen der Zentrifugal- und der Lorentzkraft $\omega_c = eB/m^*$ folgt. Die gleichzeitige Wirkung des elektrischen Wechselfelds führt zu einer Beschleunigung der Elektronen, die dann Zykloidenbahnen durchlaufen. Ist die Kreisfrequenz der Mikrowellen gleich der Zyklotronkreisfrequenz ω_c oder ein ganzzahliges Vielfaches davon, so tritt Resonanz ein, was sich in

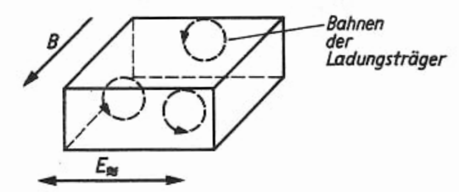


Abb. 8.2.3-5 Bestimmung der effektiven Masse mittels Zyklotronresonanz; E elektrisches Wechselfeld, $\boldsymbol{B}$ Magnetfeld

einer maximalen Energieaufnahme durch genaue Messung der Senderverlustleistung registrieren läßt. Um scharfe Resonanzen zu erhalten, muß man die Messung bei tiefen Temperaturen, meist bei der Temperatur des flüssigen Heliums, durchführen. Anderenfalls ist die mittlere freie Weglänge der Kristallelektronen zwischen zwei Stößen mit Phononen nicht groß genug. Gewöhnlich findet man mehrere effektive Massen, die verschiedenen Gruppen von Kristallelektronen bzw. verschiedenen Bahnebenen relativ zu den Gitterrichtungen entsprechen.

8.2.3.5. Defektelektronen und Löcherleitung. Enthält ein sonst leeres Band nur wenige Elektronen, so werden diese Ladungsträger entsprechend dem Verlauf der Fermi-Dirac-Verteilungsfunktion 8.2.(8) hauptsächlich die energetisch günstigsten Zustände im unteren Energiebereich des Bands einnehmen. Das Umgekehrte gilt für die unbesetzten Zustände oder Löcher in einem weitgehend gefüllten Band. Solche *Löcher* oder *Defektelektronen* werden vorzugsweise im Bereich nahe der oberen Bandkante anzutreffen sein, weil sich dann die Elektronen energetisch günstig in den tiefen Zuständen aufhalten. Man kann die Löcher im Elektronensee des Bands in gewisser Hinsicht mit Luftblasen in einer Flüssigkeit vergleichen. Quantitativ gilt für die Verteilung der Löcher im thermischen Gleichgewicht $f_{\text{Loch}} = 1 - f$, wenn f die Verteilungsfunktion 8.2.(8) bezeichnet.

Von größter Bedeutung vor allem für die Physik der Halbleiter ist nun die wohlgesicherte Tatsache, daß der Stromfluß in einem weitgehend gefüllten Band so beschrieben werden kann, als würden sich nicht die Elektronen, sondern die Löcher oder Defektelektronen bewegen. Die Löcher verhalten sich wie Ladungsträger mit positiver Ladung und einer effektiven Masse, die entgegengesetzt zu der von Kristallelektronen im oberen Bereich des Bands positiv ist.

Diese zunächst eigenartig anmutende Behauptung soll mit Hilfe der Bewegungsgleichung 8.2.(10) begründet werden. Wir nehmen an, daß sich ein einziger unbesetzter Zustand, also ein Loch, in einem vollständig gefüllten eindimensionalen Band befindet (Abb. 8.2.3-6). Unter dem Einfluß eines elektrischen Felds $\boldsymbol{E}$ muß die Wellenzahl für alle negativ geladenen Kristallelektronen im Zeitintervall dt um den gleichen Betrag $dk = -eE/\hbar$ abnehmen. Dadurch verschiebt sich die gesamte Elektronenkette, und das Loch geht in einen tieferen Zustand über, in dem es sich mit der positiven Gruppengeschwindigkeit $v_k = \hbar^{-1}(dE/dk)$ durch den Kristall bewegt. Man kann dies auch anders ausdrücken. Wie oben ausgeführt wurde (↑ 8.2.3.1.), muß die Summe der $\boldsymbol{k}$-Vektoren bzw. der Pseudoimpulse $\hbar\boldsymbol{k}$ für alle Elektronen eines vollbesetzten Bands verschwinden. Ist nun ein Zustand unbesetzt, so wird der Pseudoimpuls für ein Elektron mit der Gruppengeschwindigkeit v_k nicht durch ein gegenläufiges Elektron mit $-v_k$ kompensiert. Infolgedessen ergibt sich für den Strom eines besetzten Bands mit einem Loch der Beitrag $I = 0 - (-ev_k) = ev_k$. Das Loch bewegt sich im Feld somit wie ein positiv geladenes Teilchen. Im übrigen ist für die Löcher

auch die Energieskala umzukehren, also nach unten positiv zu rechnen. Das steht im Einklang mit der oben getroffenen Feststellung, wonach Löcher im Unterschied zu Elektronen das Bestreben haben, die energetisch höchsten Plätze im Band einzunehmen.

zu werden, ist eine Folge verschiedenartiger Wechselwirkungsprozesse. In einem idealen Gitter und ohne Wärmebewegung oder gegenseitige Beeinflussung müßten sich die Ladungsträger reibungsfrei ausbreiten, und der elektrische Widerstand von Festkörpern wäre Null.

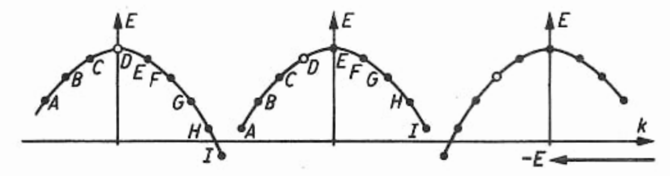


Abb. 8.2.3-6 Verschiebung eines Lochs D im $E(\mathbf{k})$-Diagramm unter Einfluß des äußeren Felds $\mathbf{E}$

Natürlich sind die Löcher keine echten physikalischen Gebilde, sondern sie bilden ein weiteres Beispiel für das in der Festkörperphysik so nützliche Konzept der Quasiteilchen. Die Zweckmäßigkeit der Konzeption fiktiver positiver Ladungsträger, die gar nichts mit den echten positiven Elektronen oder *Positronen* (s. a. 3.2.2.) zu tun haben wird vielmehr dadurch gerechtfertigt, daß es viel einfacher ist, die Bewegung weniger Löcher zu beschreiben als die Dynamik des gesamten Systems der Kristallelektronen im Band. Im übrigen darf man bei einiger Gewöhnung mit dem von Löchern getragenen Teilstrom positiver Ladungsträger genau so rechnen, wie mit dem Teilstrom negativer Elektronen in einem wenig besetzten Band.

8.2.3.6. Transportvorgänge. Die Kristallelektronen im Festkörper befinden sich auch ohne äußere Krafteinwirkungen in lebhafter Bewegung. Wenn elektrische Felder einwirken, überlagert sich dieser ungeordneten Bewegung ein gerichteter Transport von Ladungsträgern, d. h. von Elektronen und/oder Löchern, falls hierfür unbesetzte Bandzustände verfügbar sind. Für die Stromdichte $\mathbf{j}$, die in einem Feld der Stärke $\mathbf{E}$ gemessen wird, gilt bei Festkörpern im allgemeinen das Ohmsche Gesetz $\mathbf{j} = \sigma \mathbf{E}$, wobei σ die elektrische Leitfähigkeit ist. Zur atomistischen Deutung setzt man für die Leitfähigkeit

(12) $$\sigma = en\mu_n + ep\mu_p;\ \mu_n = \bar{v}_n/|\mathbf{E}|,\ \mu_p = \bar{v}_p/|\mathbf{E}|.$$

Hier bezeichnen e die Elementarladung, n bzw. p die Ladungsträgerkonzentrationen von beweglichen Elektronen (n nach „negativ") bzw. Löchern (p nach „positiv") und μ_n bzw. μ_p die sogenannten *Beweglichkeiten*. Diese zuletzt genannten Größen sind wichtige Materialkonstanten. Physikalisch bedeuten sie die mittlere Driftgeschwindigkeit $\bar{v}_n$ bzw. $\bar{v}_p$, mit der die Ladungsträger in einem Feld der Stärke 1 V/m transportiert werden. Daß die Elektronen und Löcher überhaupt eine konstante Driftgeschwindigkeit erlangen, anstatt fortgesetzt beschleunigt

Tatsächlich findet aber stets eine Streuung der Ladungsträger statt, wodurch sie nach einer mittleren freien Flugzeit τ_k, die selbst vom Wellenvektor $\mathbf{k}$ abhängt, in Streuprozessen ihre im Feld erlangte Zusatzgeschwindigkeit wieder einbüßen. Für den Zusammenhang zwischen der mittleren freien Flugzeit und der Beweglichkeit gilt $\mu = e\tau_k/m^*$; es tritt also wieder die effektive Masse auf, die für Elektronen und Löcher wesentlich verschieden sein kann. Die wichtigsten Streumechanismen, die für das Auftreten eines elektrischen Widerstands verantwortlich sind, sollen im folgenden charakterisiert werden.

8.2.3.6.1. *Elektron-Phonon-Wechselwirkung.* Dieser wichtige Effekt beruht darauf, daß die Elektronenwellen gestreut werden, wenn die strenge Gitterperiodizität durch Gitterschwingungen gestört ist. Anschaulich läßt sich der Streumechanismus durch Stoßprozesse zwischen Quasiteilchen, nämlich Kristallelektronen bzw. Löchern, und Phononen beschreiben (s. a. 8.1.2.2.). Da die thermische Bewegung bzw. die Anzahl der Phononen mit wachsender Temperatur zunimmt, ist dieser Effekt für die Abnahme der Beweglichkeit mit der Temperatur und das vor allem bei Metallen beobachtete Ansteigen des elektrischen Widerstands mit zunehmender Temperatur verntwortlich.

8.2.3.6.2. *Elektron-Elektron-Wechselwirkung.* Auch die Streuung der Ladungsträger durch Stöße untereinander begrenzt die Beweglichkeit. Allerdings ist der Effekt kleiner als man zunächst erwarten sollte, da sich die Elektronen als Folge des Pauliprinzips gewissermaßen „aus dem Wege gehen".

8.2.3.6.3. *Streuprozesse an Gitterdefekten.* Fremdatome, Versetzungen, Korngrenzen und Oberflächen behindern gleichfalls die ungestörte Ausbreitung von Ladungsträgern, wodurch die Beweglichkeit herabgesetzt wird. Aufgrund ihrer elektrischen Ladung ergibt sich ein besonders starker Einfluß für ionisierte Störstellen, was

vor allem bei Halbleitern eine wichtige Rolle spielt (↑ 8.3.2.).

Die genaue theoretische Beschreibung der elektrischen Leitfähigkeit stellt ein kompliziertes Problem dar. Teilaspekte werden in Verbindung mit Metallen und Halbleitern in späteren Abschnitten erläutert(↑8.3.2.,8.4.2.).

Die elektrische Leitung bildet nur einen, allerdings sehr wichtigen Fall von Transportvorgängen, bei denen Ladungsträger im Festkörper verschoben werden. Weitere Triebkräfte für Transportprozesse können Temperaturdifferenzen, Konzentrationsgradienten oder Inhomogenitäten der atomaren Struktur sein. Ihre theoretische Behandlung ist im Rahmen einer allgemeinen Theorie der Transportvorgänge von Elektronen und Löchern im Festkörper möglich. In diesem Buch werden als spezielle Beispiele der *Elektronenanteil zur Wärmeleitung in Metallen* (↑ 8.4.3.3.), die *thermoelektrischen und thermomagnetischen Effekte* (↑ 8.4.3.2.) und die *Diffusionsströme in Halbleitern* (↑ 8.3.3.) diskutiert.

8.2.4. Grenzen des Bändermodells – kollektive Phänomene

8.2.4.1. Leistungsfähigkeit und Grenzen des Modells. Das Bändermodell der Elektronen im Festkörper muß zweifellos als eine der bedeutendsten Errungenschaften der modernen Physik angesehen werden. Seinem Wesen nach stellt es eine Vereinigung zwischen den Erkenntnissen vom atomaren Aufbau der Kristalle und den Gesetzmäßigkeiten der Quantentheorie dar. Zu den bemerkenswerten Leistungen des Modells gehört die Deutung der elektrischen Leitungsvorgänge sowie das optische Verhalten der Festkörper. Die moderne Mikroelektronik und die Festkörperoptik sowie die beide Teilgebiete verbindende *Optoelektronik* (↑ 8.8.3.) bauen wesentlich auf den Erkenntnissen auf, die mit dem Bändermodell gewonnen wurden, aber auch das tiefergehende Verständnis mechanischer Werkstoffeigenschaften wird hiervon befruchtet. Wie jedes physikalische Modell, hat aber auch das Bändermodell Grenzen seiner Anwendbarkeit. Zunächst wurde die Theorie für den weitgehend idealen Kristall entwickelt. Selbst die perfektesten Einkristalle von festen Stoffen enthalten aber viele Baufehler, und häufig werden die beobachteten elektrischen und optischen Eigenschaften sehr wesentlich von diesen Defekten bestimmt. Glücklicherweise läßt sich das Bändermodell erweitern, indem man diskrete Elektronenzustände und die entsprechenden lokalisierten Elektroneneigenfunktionen, die solchen Defekten zuzuordnen sind und vielfach in die verbotenen Energiebereiche fallen, nachträglich in das Modell einbaut. Davon wurde verschiedentlich schon Gebrauch gemacht. Ernster ist die Situation bei den nichtkristallinen oder amorphen Festkörpern, bei denen keine periodische Fernordnung der Atome vorliegt (s. a. 6.4.6.). Hierzu gehören z. B. die amorphen Metalle oder Metgläser, etwa Fe/B oder Legierungen der Zusammensetzung $Fe_{40}Ni_{40} \cdot P_{14}B_6$, welche durch sehr rasches Abkühlen von Schmelzen erhalten werden, indem ein feiner Strahl auf rotierende, gekühlte Kupferwalzen aufgespritzt wird („splatcooling"-Technik). Auch amorphes Silicium, das durch Abscheidung aus der Gasphase bei niedrigen Temperaturen unter Einbau von stabilisierendem Wasserstoff (a-Si : H) in Form dünner Schichten gewonnen werden kann, hat neuerdings als aussichtsreiches Halbleitermaterial, besonders für Sonnenbatterien, großes Interesse gefunden. Um die elektrischen oder optischen Eigenschaften solcher neuen Werkstoffe zu beschreiben, reicht die Begründung des Bändermodells offensichtlich nicht aus. Trotzdem hat es den Anschein, daß man wesentliche Erkenntnisse des Modells, wenngleich mit einigen Modifikationen, auch auf amorphe Festkörper anwenden kann. Warum dies möglich ist, bedarf allerdings im einzelnen noch der Klärung. Hinsichtlich des Elektronentransports in stark gestörten Festkörpern, vor allem in Isolatoren, sind gleichfalls Ergänzungen nötig. Anstelle der Konzeption sich fortpflanzender Elektronenwellen, die Streuprozessen unterliegen, muß man dann aufeinanderfolgende Sprünge lokalisierter Elektronen über dazwischenliegende Barrieren betrachten. Dieser von Mott eingeführte Mechanismus wird als *Hoppingleitung* bezeichnet (s. a. 8.6.3.2.).

8.2.4.2. Kollektive Elektronenphänomene. Prinzipiell wird das Bändermodell auch dadurch begrenzt, daß es sich um eine Einelektronentheorie handelt, bei der nur die Zustände und Eigenfunktionen einzelner, aus dem Kollektiv als repräsentant herausgegriffener Elektronen betrachtet werden. Unberücksichtigt bleiben dabei naturgemäß solche Erscheinungen, bei denen die Wechselwirkung der Kristallelektronen untereinander eine wesentliche Rolle spielt. Die neuere Forschung hat jedoch ergeben, daß wichtige Phänomene in Festkörpern nur verstanden werden können, wenn man kollektive Zustände der Kristallelektronen einbezieht. Beispiele hierfür bilden Polaronen, Exzitonen, Cooperpaare und Plasmonen.

8.2.4.2.1. *Polaronen.* Elektronen, die sich durch den Festkörper bewegen, rufen aufgrund ihrer elektrischen Ladung in der Umgebung eine Polarisation hervor. Dabei werden die gleichnamig geladenen Elektronen etwas zurückgedrängt, die benachbarten positiv geladenen Atomrümpfe dagegen etwas angezogen. Wegen der geringeren Trägheit werden allerdings in erster Näherung nur die Elektronen verschoben. Die

so aufgebaute Polarisationswolke wird vom Elektron mitgeführt und bewirkt eine Vergrößerung der effektiven Masse. Dies ist vor allem in isolierenden Ionenkristallen ein gut meßbarer Effekt (s. a. 8.6.3.). Bei der Ausbreitung durch den Kristall verhält sich das Elektron dann so, als wäre es von einer Wolke optischer Phononen umgeben, die mit dem Gitter ausgetauscht werden. Der so geschaffene energetische Anregungszustand wird in der Theorie als Quasiteilchen behandelt und Polaron genannt.

8.2.4.2.2. *Exzitonen.* Wenn in einen Kristall genügend energiereiche Lichtquanten eingestrahlt werden, werden Elektronen aus dem Valenzband unter Überspringen der Energielücke in das Leitungsband angeregt, und es bilden sich *Elektron-Loch-Paare.* Infolge der Anziehung zwischen den negativ geladenen Elektronen und den quasipositiv geladenen Löchern können diese Teilchen unter geeigneten Bedingungen beisammenbleiben, indem sie sich wechselseitig umkreisen und gemeinsam durch das Gitter bewegen. Energetisch entsprechen diese Exzitonen diskreten Anregungszuständen, deren Energie etwas kleiner ist als der Bandabstand. Mit optischen Mitteln konnten diese Quasiteilchen nachgewiesen werden, und der stromlose Transport von Anregungsenergie durch Exzitonen vermag verschiedenartige, sonst schwer verständliche Effekte im Ergebnis einer Lichteinstrahlung in Festkörpern zu erklären (s. a. 8.8.4.). Durch Einwirkung von Laserimpulsen hoher Leistungsdichte auf Halbleiter bei niedrigen Temperaturen (unterhalb 80 K in CdS bzw. 20 K in Si) konnte die Ausbildung neuartiger kollektiver Anregungszustände realisiert werden. Es bilden sich ein Exzitonengas relativ geringer Dichte sowie „Tropfen" einer Art „Quantenflüssigkeit" aus eng gepackten Elektronen und Löchern, die metallische Leitfähigkeit und ungewöhnliche optische Eigenschaften aufweisen. Die Tropfen dieser *Elektron-Loch-Flüssigkeit* zerfallen nach dem Abschalten des Laserimpulses mit Halbwertszeiten von 10^{-9} bis 10^{-5} s.

8.2.4.2.3. *Cooperpaare.* Bei sehr niedrigen Temperaturen können sich Elektronen in Festkörpern, hauptsächlich in Metallen, paarweise bei entgegengerichtetem Spin zusammenschließen. Dieser überraschende Effekt, der auf einer Wechselwirkung unter Vermittlung der Phononen des Gitters beruht, hat den Schlüssel zum atomistischen Verständnis der *Supraleitung* geliefert (↑ 8.5.3.).

8.2.4.2.4. *Plasmonen.* Läßt man einen Strahl von Elektronen einheitlicher Energie durch dünne Schichten fester Stoffe hindurchtreten und mißt den Energieverlust der durchgelassenen Teilchen mit großer Genauigkeit, so findet man diskrete Werte, die Vielfachen von $\hbar\omega_p$ entsprechen. Mit ω_p werden dabei die Plasmonenfrequenzen bezeichnet, die je nach Art und Struktur des festen Stoffs charakteristische Werte besitzen. Der interessante Effekt ist als eine kollektive Anregung des Elektronensystems relativ zum Gitter der Atomrümpfe zu deuten. In der Nähe des bestrahlten Gebiets schwingt das System der Kristallelektronen oder Teile davon relativ zum Gitter hin und her. Diese Anregungszustände, die gequantelt sind, werden als Plasmonen bezeichnet, da gewisse Ähnlichkeiten zu Schwingungen in Gasplasmen bestehen (s. a. 9.1.2.). Die Methode, die auch mit reflektierten Elektronen an Festkörperoberflächen angewendet werden kann, hat als *Elektronen-Energie-Verlust-Analyse* (EEVA) zunehmende Bedeutung für die Untersuchung der Struktur von Festkörpern erlangt (Abb. 8.2.4-1).

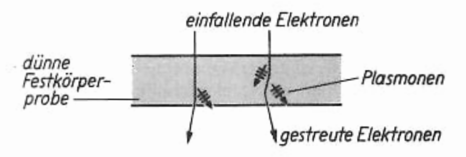


Abb. 8.2.4-1 Prinzip der Elektronen-Energie-Verlust-Analyse

8.3. Halbleiter

8.3.1. Merkmale und Arten der Halbleiter

Als Halbleiter werden Stoffe gekennzeichnet, die hinsichtlich ihrer elektrischen Leitfähigkeit eine Mittelstellung zwischen den metallischen Leitern und den Isolierstoffen oder Dielektrika einnehmen. Die *spezifische Leitfähigkeit von Halbleitern* umfaßt den Bereich 10^{-7} bis $10^{4}\ \Omega^{-1}\mathrm{m}^{-1}$, jedoch wird die Abgrenzung zu den Isolierstoffen, die nicht prinzipieller Natur tsi, häufig nach der Zweckmäßigkeit vorgenommen. Obwohl erste Untersuchungen und auch Anwendungen der Halbleitereigenschaften bis in das vergangene Jahrhundert zurückreichen, konnte die stürmische Entwicklung der Halbleitertechnik etwa seit dem Beginn der 50er Jahre erst einsetzen, nachdem seitens der Grundlagenforschung zwei wichtige Voraussetzungen geschaffen worden waren: die Fertigkeit zur Herstellung weitgehend perfekter Einkristalle höchster Reinheit und das zum Verständnis der Halbleiter unentbehrliche *Bändermodell der Kristallelektronen* (↑ 8.2.).

Nach den Erkenntnissen des Bändermodells lassen sich Halbleiter als Festkörper definieren, bei denen der Bandabstand E_g zwischen dem bei 0 K vollständig gefüllten Valenzband und dem darüberliegenden Leitungsband höchstens 2,5 bis 3 eV beträgt, wobei das Ferminiveau E_F in der verbotenen Zone zwischen beiden Bän-

dern liegt. Damit lassen sich auch die folgenden Merkmale der Halbleiter atomistisch deuten.
Die Leitfähigkeit halbleitender Stoffe hängt stark von der Temperatur ab; sie ist bei tiefen Temperaturen sehr klein und nimmt mit wachsender Temperatur im allgemeinen exponentiell zu. Allerdings kann die Leitfähigkeitskurve bei gezielt oder ungewollt verunreinigten Halbleitern Temperaturbereiche mit nahezu konstanter oder sogar abnehmender Leitfähigkeit aufweisen.
Völlig reine Halbleiter leiten den Strom relativ schlecht, aber schon geringe Spuren von Verunreinigungen können die Leitfähigkeit um Größenordnungen erhöhen. Auf der gezielten Beeinflussung der Leitfähigkeit durch Einlagerung von Fremdatomen, die man als *Dotierung* bezeichnet, beruht letztlich der Einsatz von Halbleitern in der modernen Elektronik.
Auch Licht oder allgemeiner elektromagnetische Strahlen ausreichender Quantenenergie, ferner Teilchenstrahlen sowie elektrische und magnetische Felder können die Leitfähigkeit von Halbleitern erheblich verändern. Diese Effekte werden in verschiedenartigen *Halbleiterbauelementen* (Tafel 44/45) sowie für *Solarzellen* (↑ 8.8.3.2.) genutzt. Die vorstehend aufgeführten Merkmale lassen sich qualitativ leicht verstehen, wenn man davon ausgeht, daß die für den Stromtransport erforderlichen Ladungsträger, nämlich Elektronen in dem weitgehend leeren Leitungsband und Löcher im gefüllten Valenzband, durch verschiedene Ursachen erzeugt werden können (Abb. 8.3.1-1). Durch thermische Anregung (↑ 8.2.2.2.) entstehen Elektronen und Löcher paarweise, indem Elektronen unter Zurücklassung von Löchern in das Leitungsband angehoben werden. Wenn nur thermische Anregung vorliegt und auch keine Ladungsträger von Fremdatomen abgegeben werden, spricht man von *Eigenleitung*. Dagegen herrscht *Störstellenleitung* vor, wenn der Stromtransport vorwiegend von Ladungsträgern getragen wird, die von Fremdatomen oder Störstellen hervorgerufen werden. Überwiegen dabei die Elektronen im Leitungsband, also die negativen Ladungsträger, so wird der Halbleiter als *n*-leitend bezeichnet. Im entgegengesetzten Fall der *p*-Leitung wird der Strom vorwiegend von den effektiv positiven *Löchern* (↑ 8.2.3.5.) getragen, die zurückbleiben, wenn Elektronen aus dem Valenzband von Störstellen eingefangen werden. Homogene Halbleiter besitzen eine einheitliche Zusammensetzung bei gleichmäßiger Verteilung der Fremdatome. Für die Halbleitertechnik und die *Mikroelektronik* sehr wichtig sind inhomogene halbleitende Strukturen, bei denen Schichten oder auch sehr kleine Bereiche mit *p*- bzw. *n*-Leitung aneinander grenzen, was durch gezielte Dotierung von Kristallen des gleichen Grundmaterials mit verschiedenartigen Fremdatomen erreicht werden kann. Bevor wir auf diese Probleme näher eingehen, soll eine Übersicht über die verschiedenen Arten von Halbleitern gegeben werden (Tab. 8.3.1-2). Halbleitende Elemente finden sich in den Hauptgruppen des mittleren und rechten Teils des Periodensystems der Elemente. Beispiele sind Bor, Silicium, Germanium, Arsen, Selen, Tellur. In einigen Fällen, vor allem bei Kohlenstoff und Zinn, ist das Auftreten halbleitender Eigenschaften an bestimmte Festkörperphasen gebunden, während in anderen Modifikationen isolierende bzw. metallische Eigenschaften auftreten. Von den anorganischen Verbindungshalbleitern sind bisher vorwiegend binäre und ternäre Kombinationen aus zwei bzw. drei Elementen erforscht worden. Bei den $A^{III}B^{V}$-Verbindungen, von denen Galliumarsenid GaAs ein wichtiger Vertreter ist, gehört eine Komponente zur III. und die andere zur V. Hauptgruppe des Periodensystems. Analog bildet das für fotoelektrische Schichten bedeutsame Cadmiumsulfid CdS ein Beispiel für eine $A^{II}B^{VI}$-Verbindung. Halbleiterverhalten wird auch bei Gläsern und sogar bei einigen Schmelzen beobachtet. Zu erwähnen sind die sogenannten Chalkogenidgläser, die aus Elementkombinationen wie As/Se/Te oder As/Te/Si/Ge aufgebaut sind. Schließlich haben in neuerer Zeit auch halbleitende organische Molekülkristalle zunehmendes Interesse gefunden. Solche *orga-*

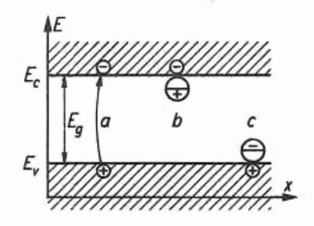


Abb. 8.3.1-1 Freisetzung von Ladungsträgern in Halbleitern durch Band-Band-Übergang (*a*) bzw. Ionisierung von Störstellen (*b*, *c*); E_g Bandabstand, E_c, E_v Bandkanten

Tab. 8.3.1-2 Eigenschaften einiger Halbleiter; E_g Bandabstand in eV bei 300 K, ε Dielektrizitätskonstante, μ_n, μ_p Beweglichkeit der Elektronen bzw. der Defektelektronen in cm^2/Vs

Stoff	E_g	ε	μ_n	μ_p
Si	1,09	12	1500	500
Ge	0,66	16	3800	1800
SiC	2,8	6,7	100	20
CaAs	1,4	11	8000	400
InSb	0,17	16	70000	4000
GdS	2,4	11,6	200	20

nischen Halbleiter besitzen meist abwechselnd Einfach- und Doppelbindungen, sogenannte konjugierte Bindungen, sowie teilweise organisch gebundene Metallatome. Beispiele sind die Ringkohlenwasserstoffe Anthracen und Stilben, Blei-Phthalocyanin oder die hochleitfähigen TCNQ- (Tetracyanochinodimethan-) Komplexe.
Trotz der großen Vielfalt halbleitender Stoffe dominieren bei technischen Anwendungen aus physikalischen und technologischen Gründen relativ wenige Halbleitermaterialien. In der eigentlichen Halbleitertechnik und Mikroelektronik ist dies vor allem das *Silicium*, das mit einem Bandabstand $E_g = 1{,}12$ eV bei 300 K, hoher chemischer und struktureller Beständigkeit sowie den ausgefeilten Methoden zur Reindarstellung und Züchtung großer Einkristalle nahezu optimale Eigenschaften verkörpert. Daneben besitzen vor allem Germanium und Galliumarsenid wegen der größeren Ladungsträgerbeweglichkeit und der damit zusammenhängenden größeren Schaltgeschwindigkeit von Bauelementen eine gewisse Bedeutung. Allerdings ist nicht zu übersehen, daß die spezifischen Anforderungen der Optoelektronik (s. a. 8.8.3.) einen wachsenden Einsatz von Halbleitern mit einem Bandabstand um 2 eV, wie Galliumphosphid GaP, bedingen. Auf der anderen Seite werden für Sensoren zum Nachweis infraroter Strahlen „schmalbandige" Halbleiter mit kleinem E_g wie Quecksilber-Blei-Tellurid benötigt. Überhaupt zwingt der wachsende Bedarf, besonders in Verbindung mit der Automatisierung und Roboterentwicklung, an Sensoren, d. h. künstlichen Sinnesorganen für optische, akustische, thermische, magnetische oder elektrische Einflüsse, zu einer intensiveren Entwicklung von Spezialbauelementen, für die auch spezifische Materialien benötigt werden. Schließlich sollte auch nicht übersehen werden, daß Halbleitereffekte nicht nur in der Elektronik, sondern beispielsweise auch in der Solarenergietechnik (s. a. 8.8.3.) oder in der Elektrofotografie, darunter bei Schnellkopierverfahren (s. a. 8.8.4.) eine wichtige Rolle spielen. Dies alles ist Anlaß, die halbleitenden Eigenschaften der verschiedensten Stoffe grundlagenmäßig sorgfältig zu erforschen.

8.3.2. Eigenleitung

8.3.2.1. Berechnung der Ladungsträgerdichten. Bei reiner Eigenleitung werden die Ladungsträger ausschließlich durch thermische Anregung von Elektronen aus dem Valenzband in das Leitungsband gebildet. Da die Ladungsträger dabei paarweise entstehen, muß $n = p$ gelten, d. h. die Konzentration der Elektronen im Leitungsband n muß gleich der Konzentration der Löcher im Valenzband p sein. Durch Integration des Ausdrucks für die Elektronendichte [8.2.(9)] und des analogen Terms für die Löcherdichte lassen sich diese Konzentrationen mit Hilfe der Zustandsdichtefunktion $D(E)$ und der Fermi-Dirac-Verteilungsfunktion $f(E)$ berechnen (↑ 8.2.2.3.). Dies führt auf die Integrale

$$(1) \quad n = \int_{E_c}^{E_c'} D(E)\, f(E)\, dE;$$

$$p = \int_{E_v'}^{E_v} D(E)\,[1 - f(E)]\, dE.$$

Der Verlauf der einzelnen Funktionen sowie der Integranden ist in Abb. 8.3.2-1 dargestellt; das

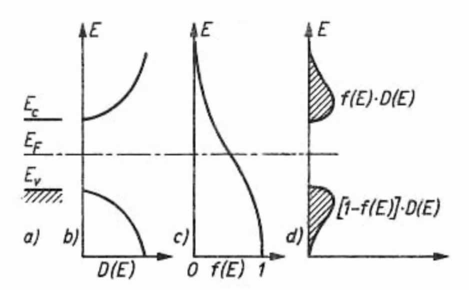


Abb. 8.3.2-1 Berechnung der Trägerdichten bei Eigenleistung; a) Bändermodell, b) Zustandsdichte $D(E)$, c) Verteilungsfunktion $f(E)$, d) Elektronenkonzentration in den Bändern

Ergebnis entspricht den schraffierten Flächen im rechten Teilbild. Die analytische Auswertung wird erleichtert, da die Integranden wegen des exponentiellen Abfalls von $f(E)$ bzw. $[l - f(E)]$ nur in der Nähe der Bandkanten E_c bzw. E_v von Null verschieden sind. Infolgedessen können die Integrationsgrenzen E_c' und E_v' durch $+\infty$ bzw. $-\infty$ ersetzt werden. Für nicht zu kleine Bandabstände, d. h. $(E_c - E_F) \gg k_B T$ und $(E_F - E_v) \gg k_B T$, führt die Rechnung auf

$$(2) \quad n = N_c \exp[-(E_c - E_F)/k_B T];$$

$$p = N_v \exp[-(E_F - E_v/k_B T].$$

Die hier eingeführten Größen N_c und N_v sind für einen gegebenen Halbleiter näherungsweise Stoffkonstanten; sie heißen *effektive Zustandsdichten* des Valenz- bzw. des Leitungsbands.
Das Ferminiveau E_F liegt bei Eigenhalbleitern nur für $T = 0$ K genau in der Mitte zwischen den Bandkanten E_c und E_v. Bei höheren Temperaturen ergeben sich Abweichungen von der Bandmitte, falls die effektiven Massen von Elektronen und Löchern nicht übereinstimmen, und zwar nähert sich das Ferminiveau dann dem Band mit der kleineren effektiven Masse der Ladungsträger.

8.3.2.2. Massenwirkungsgesetz der Ladungsträger. Ein grundlegendes Gesetz der Halbleiterphysik erhält man durch Bildung des Produkts der beiden Ausdrücke von 8.3.(2):

$$(3) \quad np = n_i^2 = N_c N_v \exp[-(E_c - E_v)/k_B T].$$

Da die Fermienergie E_F in dieser Gleichung nicht mehr auftritt, hängt die Größe np nur noch von der Temperatur und von dem für einen gegebenen Stoff konstanten Bandabstand $E_g = E_c - E_v$ ab. Das Produkt der Ladungsträgerkonzentrationen ist demnach für einen bestimmten Halbleiter bei gegebener Temperatur konstant. Die Wurzel aus der Konstanten wird als *Inversionskonzentration* n_i bezeichnet. Im Falle der Eigenleitung ist $n = p = n_i$. Bei Zimmertemperatur beträgt $n_i^2 = 5{,}76 \cdot 10^{32}\ \text{m}^{-6}$ für Germanium und $n_i^2 = 2{,}25 \cdot 10^{26}\ \text{m}^{-6}$ für Silicium. In eigenleitendem Ge bzw. Si sind somit bei 20 °C pro Kubikzentimeter größenordnungsmäßig 10^{13} bzw. 10^{10} Ladungsträger vorhanden. Da 1 cm³ rund 10^{22} Atome enthält, kommt daher erst auf 10^9 Ge- oder 10^{12} Si-Atome ein bewegliches Elektron und ein Loch. Allerdings steigen die Ladungsträgerkonzentrationen bei Eigenleitung mit zunehmender Temperatur exponentiell an. Beim Vergleich verschiedener Materialien ist zu berücksichtigen, daß die Anzahl der Ladungsträger gleichfalls exponentiell vom Bandabstand abhängt, woraus sich der mehrere Größenordnungen betragende Unterschied in den Ladungsträgerkonzentrationen von Ge und Si erklärt. Da bei der Ableitung von Gleichung 8.3.(3) nur allgemeine Gesetzmäßigkeiten des Bändermodells verwendet und keine Aussagen über die Herkunft der Ladungsträger getroffen wurden, muß diese wichtige Beziehung auch außerhalb des Bereichs der Eigenleitung erfüllt sein. Die Gleichung sagt dann aus, daß jede Erhöhung der Konzentration einer der beiden Ladungsträgerarten mit einem entsprechenden Rückgang der Konzentration der anderen Ladungsträgerart verbunden ist. Wegen der Analogie zum Massenwirkungsgesetz bei bimolekularen chemischen Reaktionen (s. a. 7.1.1.) wird die Gleichung 8.3.(3) als *Massenwirkungsgesetz der Ladungsträger* in Halbleitern bezeichnet.

8.3.3. Störstellenleitung

8.3.3.1. Dotierung von Halbleitern. Die Erforschung des elektrischen Verhaltens von Halbleitern wurde in der Vergangenheit durch die Tatsache erschwert, daß Fremdatome, die in das Gitter eines Halbleiterkristalls eingebaut sind, Elektronen an das Leitungsband abgeben oder umgekehrt Löcher durch Aufnahme von Elektronen aus dem Valenzband bilden können. Diese von Fremdatomen oder Störstellen hervorgerufene Störstellenleitung kann die Eigenleitung völlig überdecken und die Leitfähigkeit um Größenordnungen erhöhen. Liefert beispielsweise jede Störstelle ein Elektron an das Leitungsband, so reicht eine Konzentration von einem Fremdatom auf 10^8 Ge- oder 10^{11} Si-Atome aus, um die Elektronenkonzentration im Vergleich zur Eigenleitung um eine Größenordnung zu steigern. Nach dem Massenwirkungsgesetz muß dann die Konzentration der Löcher um eine Größenordnung abnehmen, damit das Produkt np (↑ 8.3.2.2.) konstant bleibt. Das wirkt sich indessen nur wenig auf die Summe $n + p$ der Ladungsträgerkonzentrationen aus, die bei Störstellenleitung im wesentlichen durch die im Überschuß vorliegenden *Majoritätsträger* bestimmt wird. Die Leitfähigkeit nimmt daher als Folge der Einlagerung von elektronenspendenden Störstellen zu. Die mit kleinerer Konzentration noch vorhandenen Ladungsträger der anderen Art, im betrachteten Beispiel also die Löcher, werden *Minoritätsträger* genannt. Fremdatome, die Elektronen an das Leitungsband abgeben, heißen *Donatoren*. Im Bänderschema werden sie durch diskrete Niveaus dicht unter der Kante E_c des Leitungsbands veranschaulicht (Abb. 8.3.3-1). Beispiele bilden Phosphoratome in Silicium oder Antimonatome in Germanium, die anstelle der vier Bindungselektronen der Halbleiteratome fünf Außenelektronen mitbringen. Umgekehrt verhalten sich Fremdatome mit nur drei Außenelektronen wie Bor oder Indium. Sie neigen dazu, das fehlende Elektron aus dem Valenzband des Wirtsgitters aufzunehmen und werden *Akzeptoren* genannt.

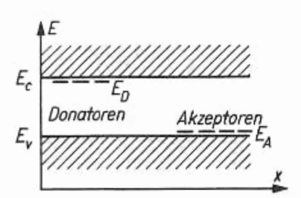


Abb. 8.3.3-1 Lage typischer Donator- bzw. Akzeptorterme im Bändermodell

Allgemein bewirkt die Einlagerung von Donatoren, daß die Elektronen Majoritätsträger werden und der Halbleiter n-Leitung zeigt. Umgekehrt führt der Einbau von Akzeptoren zum Auftreten von p-Leitung mit den Löchern als Majoritätsträger. Der Einfluß nur einer Störstellenart hat in jedem Falle eine Erhöhung der gesamten Ladungsträgerkonzentration $n + p$ zur Folge, wodurch die Leitfähigkeit um Größenordnungen steigen kann. Enthält das Halbleitermaterial jedoch sowohl Donatoren als auch Akzeptoren, so kann bei gleicher Konzentration und

gleicher Wahrscheinlichkeit zur Aufnahme bzw. Abgabe von Elektronen eine *Störstellenkompensation* eintreten. Im Idealfall, der praktisch aber kaum zu verwirklichen ist, wird der Halbleiter dann wieder eigenleitend. Bei dem meist vorliegenden Überschuß einer Art von Störstellen entscheidet die vorherrschende Art den

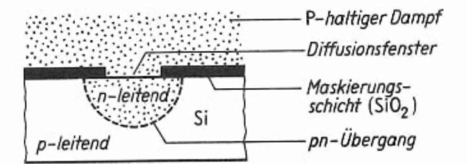


Abb. 8.3.3-2 Erzeugung eines *pn*-Übergangs in Silicium durch Umdotieren mit Hilfe der Diffusionstechnik

Leitungscharakter. Es leuchtet ein, daß der drastische Einfluß von spurenhaften Verunreinigungen auf den Leitungstyp sowie die Leitfähigkeit extreme Forderungen an die Reinheit definierter Halbleitermaterialien stellt. Ausgehend von sehr reinem Material läßt sich dagegen der Leitungstyp durch die gezielte Einlagerung von Donatoren oder Akzeptoren in das Gitter in gewünschter Weise verändern und die Leitfähigkeit in weiten Grenzen verschieben. Dieser für die Halbleitertechnologie grundlegende Prozeß wird als *Dotierung* bezeichnet, und die absichtlichen Fremdstoffzusätze heißen *Dotanten*. Um Halbleiterkristalle zu dotieren, kann man kleine Anteile entsprechender Elemente, bei Silicium beispielsweise Bor oder Phosphor, zur Schmelze vor dem Ziehen der Einkristalle zusetzen. Läßt man Störstellenatome des entgegengesetzten Typs aus der Gasphase bei hohen Temperaturen in die oberflächennahe Zone des Kristalls eindiffundieren, so kommt es in diesem Bereich zunächst zur Störstellenkompensation und danach zur Umdotierung. Auf diese Weise läßt sich ein *pn*-Übergang mit aneinandergrenzenden *p*- und *n*-leitenden Bereichen schaffen (Abb. 8.3.3-2, s. a. 8.3.6.1.). Dieser Vorgang kann mehrfach wiederholt werden. Sorgt man dabei durch geeignete Maskierung der Halbleiteroberfläche dafür, daß der Dotant nur an bestimmten Stellen eindiffundieren kann, so lassen sich zahlreiche, unter Umständen viele tausende, von umdotierten inselartigen Bereichen in ein winziges Scheibchen eines Si-Kristalls einarbeiten, wie bei der Behandlung der *Mikroelektronik* noch ausgeführt wird (↑ 8.3.6.). Eine weitere Möglichkeit zur Dotierung bietet die *Ionenimplantation*, die neuerdings in zunehmendem Umfang technisch eingesetzt wird. Man beschießt dabei die Halbleiterproben mit energiereichen Ionen des gewünschten Dotierungsmaterials. Zu diesem Zweck wurden spezielle Ionenquellen und leistungsfähige Beschleuniger bis zu Ionenenergien von einigen 10^2 keV entwickelt. Die Ionen dringen, vor allem längs relativ freier Kanäle, in das Kristallgitter ein, wo sie abgebremst und schließlich als Störstellen eingebaut werden. Allerdings werden dabei durch Verschiebung regulärer Gitterbausteine durch Ionenstöße relativ viele Kristalldefekte geschaffen, die mittels einer geeigneten thermischen Behandlung weitgehend ausgeheilt werden müssen. Schließlich sei als moderne Methode zum Einbau von gezielten Störstellen noch die *Neutronendotierung* erwähnt. Bei diesem originellen Verfahren, das ebenfalls technisch genutzt wird, wandelt man einen kleinen Bruchteil der Atome des Halbleiters durch Kernreaktionen mittels Neutronenbestrahlung in einem Atomkernreaktor (s. a. 3.4.2. und 3.5.2.) in Atome des Dotanten um. So bildet sich aus dem in Si zu reichlich 3 % enthaltenen Isotop $^{30}_{14}Si$ nach einer (n,γ)-Reaktion radioaktives $^{31}_{14}Si$, das durch einen β^--Zerfall mit einer Halbwertszeit von 2,62 h in stabilen $^{31}_{15}P$ übergeht.

8.3.3.2. Substitutionsstörstellen als Donatoren bzw. Akzeptoren. Der Einbau von Fremdatomen in Halbleiterkristalle und ihre Rolle als Donatoren bzw. Akzeptoren soll im folgenden etwas näher betrachtet werden. Am übersichtlichsten liegen die Verhältnisse, wenn die Fremdatome auf normalen Gitterplätzen eingebaut sind und entweder ein Außenelektron mehr oder eines weniger besitzen als die Atome des Wirtsgitters. Da die Fremdatome dann normale Kristallbausteine vertreten, bezeichnet man sie als *Substitutionsstörstellen*. Im Rahmen einer allgemeinen Klassifikation der *Realstruktur von Kristallen* (↑ 6.4.7.) bilden die Substitutionsstörstellen einen Spezialfall der Punktdefekte oder der nulldimensionalen Fehlordnung. Da zum Unterschied von einfachen Gitterlücken oder Wirtsgitteratomen auf Zwischengitterplätzen Fremdatome anderer Elemente beteiligt sind, spricht man auch von chemischer Fehlordnung. Die Kristalle des Wirtsgitters können im allgemeinen solche Fremdatome nur in geringer Konzentration aufnehmen, ohne daß starke Gitterverzerrungen resultieren, die schließlich die Entstehung gröberer Gitterdefekte, wie Versetzungen oder Ausscheidungen, zur Folge hätten. Am leichtesten werden naturgemäß Fremdatome eingebaut, die sich in der Größe nur wenig von denen des Grundmaterials unterscheiden. Das ist vor allem für Atome der Elemente erfüllt, die zu Nachbarn oder darüber- bzw. darunterliegenden Reihen im Periodensystem der Elemente gehören. Beispiele seien die auch technisch wichtigsten Fälle des Einbaus von Phosphor bzw. Bor in Silicium. Abb. 8.3.3-3 veranschaulicht den Einbau in schematischer Weise, wobei die räumliche Diamantstruktur

der Siliciumkristalle „plattgedrückt" wurde. Um die Energie abzuschätzen, die zur Ionisierung eines Phosphoratoms gemäß der Reaktion $P \rightarrow P^+ + \ominus$ unter Abgabe des Elektrons $\ominus$ an das Leitungsband notwendig ist, faßt man

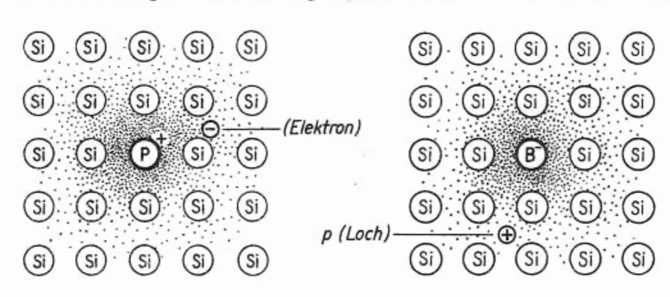


Abb. 8.3.3-3 Donator- bzw. Akzeptorstörstelle in Silicium (schematisch)

die ionisierte Störstelle P^+ näherungsweise als einfach positiv geladenes, punktförmiges Zentrum auf. Um diesen Rumpf soll sich das überzählige 5. Außenelektron in ähnlicher Weise bewegen wie das Elektron im Bohrschen Modell des Wasserstoffatoms (↑ 4.2.2.). Für das Elektron im Grundzustand des Wasserstoffatoms beträgt der Bahnradius $a_0 = \varepsilon_0/\pi\hbar^2 m_e e^2$ und $E_0 = -1/2\, m_e e^4 (4\pi\varepsilon_0\hbar)^{-2}$ ist die Bindungsenergie. Nun bewegt sich das Hüllenelektron im Falle der P-Störstelle nicht wie beim Wasserstoffatom im Vakuum, sondern im polarisierbaren Medium des Gitters und der Kristallelektronen. In pauschaler Weise kann man dem Rechnung tragen, indem in den obigen Beziehungen an Stelle von ε_0 das Produkt $\varepsilon_0\varepsilon$ mit der relativen Dielektrizitätskonstanten ε des Festkörpers gesetzt wird. Für Silicium ist $\varepsilon = 12$. Außerdem ist die Elektronenmasse m_e durch die effektive Masse m^* zu ersetzen, die für Leitungselektronen in Si im Mittel etwa 0,31 m_e beträgt. Daraus folgt eine fast 40fache Vergrößerung des Bahnradius gegenüber dem Wasserstoffmodell. Das Elektron bewegt sich daher relativ weit außerhalb der eigentlichen Störstelle, wodurch die Benutzung der Dielektrizitätskonstanten des makroskopischen Materials nachträglich gerechtfertigt wird. Die Bindungsenergie verringert sich um einen Faktor $\varepsilon^2 m_e/m^*$, woraus nach Einsetzen der Zahlenwerte der Betrag $E_B = E_c - E_D = -0{,}03$ eV folgt. Nach Messungen liegen die Donatorniveaus E_D für P in Si um etwa 0,045 eV unter der Kante E_c des Leitungsbands (Abb. 8.3.3-4).

In ganz analoger Weise läßt sich die Energie abschätzen, mit der ein Loch an den negativen Rumpf einer B^--Störstelle „gebunden" ist, auch wenn dieses Modell recht ungewöhnlich erscheint. Man findet für die Lage der Akzeptorniveaus über der Kante E_v des Valenzbands $E_B = E_A - E_v = 0{,}034$ eV gegenüber Meßwerten von gleichfalls 0,45 eV.

Insgesamt zeigen diese Abschätzungen trotz der mangelhaften quantitativen Übereinstimmung mit den Meßdaten, warum Substitutionsstörstellen schon mit geringem Energieaufwand in den ionisierten Zustand übergehen können, wobei sie im Falle von Donatoren Elektronen an das Leitungsband und im Falle von Akzeptoren Löcher an das Valenzband des Halbleiters abgeben.

8.3.3.3. Weitere Störterme. Außer den erwünschten flachen Störstellen in der Nähe der Bandkanten finden sich in Halbleitern stets auch lokalisierte Elektronenniveaus in Form *tiefliegender Störterme* innerhalb der verbotenen Zone zwischen den Bändern. Wie Abb. 8.3.3-4 für einige Beispiele von solchen Störstellen in Si ausweist, sind dafür vielfach Verunreinigungen durch Elemente verantwortlich, die weniger gut in das Kristallgitter passen. Diese Atome, die teilweise auch in den doppelt ionisierten Zustand übergehen können, sind im Kristall entweder als Substitutionsstörstellen auf normalen Gitterpositionen oder unter leichter Deformation des umgebenden Kristalls auf Zwischengitterplätzen angeordnet.

Darüber hinaus sind aber auch Elektronenniveaus von großer Bedeutung, die nicht durch

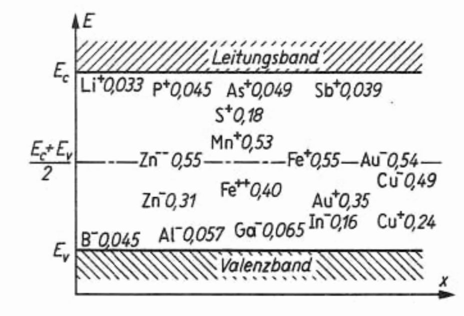


Abb. 8.3.3-4 Lage flacher und tiefliegender Störstellen in der verbotenen Zone von Silicium; die Zahlenwerte geben den Abstand zur nächstliegenden Bandkante E_c bzw. E_v in Elektronenvolt an

Fremdatome, sondern durch andere Gitterdefekte, insbesondere durch Versetzungen, Korngrenzen oder Oberflächen hervorgerufen werden. Alle tiefliegenden Störterme tragen nur wenig zur Vergrößerung der Ladungsträgerkonzentrationen bei, weil die zu ihrer Ionisierung erforderliche Energie relativ groß ist und teilweise eine Kompensation erfolgt. Dafür machen sie sich beim Betrieb von Halbleiterbauelementen in anderer Weise unangenehm bemerkbar. Einmal wirken sie als Streuzentrum, durch die die Bewegung der Ladungsträger behindert wird (s. a. 8.4.1.), und zum anderen begünstigen sie eine rasche *Rekombination* zusätzlicher Ladungsträger (↑ 8.5.2.). Aus diesen Gründen ist es erforderlich, für elektronische Bauelemente die Halbleitermaterialien nicht nur in höchster Reinheit, sondern auch in Form weitgehend perfekter Einkristalle einzusetzen. Eine Ausnahme bilden die dotierfähigen *amorphen Halbleiter*, insbesondere das a-Si : H (↑ 8.2.4.1.). In diesen Materialien sind die unterbrochenen Siliciumbindungen mit Wasserstoff, Fluor oder anderen Elementen abgesättigt, so daß sie nicht als Störniveaus wirksam werden.

8.3.3.4. Ionisierung der Störstellen und Lage des Ferminiveaus. Bei sehr tiefen Temperaturen sind auch im Störstellenhalbleiter nur wenige Ladungsträger vorhanden, weil die meisten Störstellen dann nicht ionisiert sind. Mit zunehmender Temperatur dissoziieren die Donatoren bzw. Akzeptoren, indem sich das Gleichgewicht zwischen den neutralen Störstellen D^* bzw. A^* und den ionisierten Zuständen D^+ bzw. A^- immer weiter nach rechts verschiebt:

$$(4) \quad D^* \rightleftharpoons D^+ + \ominus ; A^* \rightleftharpoons A^- + \oplus .$$

Dabei muß in nicht zu kleinen Bereichen die *Neutralitätsbedingung* $(n + N_{A^-}) - (p + N_{D^+}) = 0$ erfüllt sein, nach der die Summe der negativen Ladungen je Volumeneinheit gleich der Summe der positiven Ladungen sein muß, wenn sich der Halbleiter im thermodynamischen Gleichgewicht befindet. Hier bezeichnen N_{A^-}, N_{D^+} die Konzentrationen der ionisierten Akzeptoren bzw. Donatoren und n, p die Konzentrationen der Elektronen im Leitungsband bzw. der Löcher im Valenzband. Die Rechtfertigung für die Forderung nach räumlicher Neutralität ergibt sich daraus, daß schon relativ kleine Ladungsüberschüsse beträchtliche elektrische Potentiale aufbauen würden. Selbst bei Halbleitern außerhalb des thermischen Gleichgewichts können nur kleine Abweichungen auftreten, und man spricht dann von *Quasineutralität*. Ferner sind die Ladungsträgerkonzentrationen n und p unabhängig von der Herkunft der Elektronen und Löcher durch das Massenwirkungsgesetz $np = n_i^2$ miteinander verknüpft. Um die Bruchteile der Störstellen, die bei einer gegebenen Temperatur ionisiert sind, mit Hilfe der Fermi-Dirac-Statistik zu berechnen, muß man daher ein relativ kompliziertes Gleichungssystem lösen. Im Ergebnis wird der bereits ausgeführte Tatbestand erhalten, wonach der Lei-

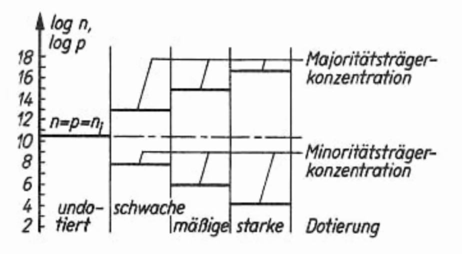


Abb. 8.3.3-5 Logarithmische Darstellung der Ladungsträgerdichten in Silicium bei $T = 100$ K für verschieden starke Dotierung

tungscharakter durch die vorherrschende Störstellenart bestimmt wird. Überwiegen beispielsweise die Donatoren, so verschiebt sich das Ferminiveau bis in die Nähe der unteren Bandkante E_c des Leitungsbands. Bei mittleren Temperaturen ist dann ein großer Teil der Donatoren ionisiert, und der Halbleiter ist n-leitend. Umgekehrt rückt das Ferminiveau bei dominierenden Akzeptoren zur Oberkante E_v des Valenzbands hin, und es tritt p-Leitung ein. Bei Störstellenkompensation durch annähernd gleiche Konzentrationen gleichwertiger Akzeptoren und Donatoren verbleibt das Ferminiveau dagegen im mittleren Gebiet der verbotenen Zone. Eine solche Annäherung an die Verhältnisse der *Eigenleitung* (↑ 8.3.2.) wird aber auch bei Überwiegen einer Störstellenart beobachtet, wenn die Temperatur genügend hohe Werte annimmt. Im Falle von Ge tritt dies etwa oberhalb 500 K, bei Si infolge des höheren Bandabstands erst oberhalb etwa 1000 K ein. Dieses Verhalten ist verständlich, weil bei genügend hohen Temperaturen die durch direkte thermische Anregung zwischen den Bändern gebildeten Ladungsträgerpaare überwiegen und den Einfluß der Störstellen zunichte machen. Eine wichtige Konsequenz ist, daß alle Halbleiterbauelemente bei hohen Temperaturen eine Verschlechterung und schließlich Einbuße ihrer Funktion zeigen, und zwar um so eher, je kleiner der Bandabstand ist. Auch aus diesem Grund wird Si in der modernen Mikroelektronik dem Ge vorgezogen. Abb. 8.3.3-5 veranschaulicht die Ladungsträgerkonzentration für verschieden starke Dotierung. Man beachte, daß in der logarithmischen Darstellung die Majoritäts- und die Minoritätsträgerkonzentrationen wegen $np = n_i^2$ symmetrisch zur Inversionskonzentration n_i liegen. Besonders übersichtliche Verhältnisse sind gegeben, wenn der Halb-

leiter nur eine Art von Störstellen, beispielsweise Donatoren der Konzentration N_D, enthält. Die Neutralitätsbedingung vereinfacht sich dann zu $n = p + N_{D^+}$. Steigt die Temperatur von tiefen Werten her allmählich an, so werden zunächst im wesentlichen nur Elektronen von den Donatoren in das Leitfähigkeitsband gelangen, denn die kleine Energiedifferenz zwischen der Unterkante des Bands und den Störstellenniveaus kann viel leichter überwunden werden als der Bandabstand. Das Valenzband bleibt infolgedessen nahezu vollständig gefüllt, und es gilt näherungsweise $n = N_{D^+}$. Solange noch neutrale Störstellen vorhanden sind, was bei Ge und Si bis in den Bereich oberhalb der Zimmertemperatur der Fall ist, spricht man von *Störstellenreserve*. Bei weiter wachsender Temperatur tritt mit $n = N_D$ *Störstellenerschöpfung* ein. Die Donatoren sind dann sämtlich ionisiert und die Ladungsträgerkonzentration ändert sich nur wenig mit der Temperatur. Erst bei noch höheren Werten von T schließt sich infolge zunehmender Anregung von Elektronen zwischen den Bändern und dem damit verbundenen Übergang zur Eigenleitung ein exponentieller Anstieg der Konzentrationen beider Ladungsträgerarten an. Abb. 8.3.3-6 zeigt den Verlauf von n als Funktion von $1/T$.

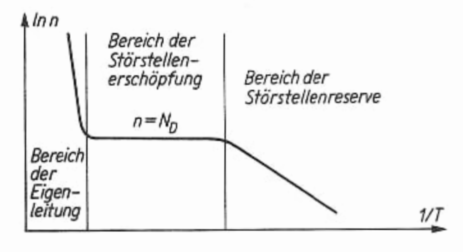


Abb. 8.3.3-6 Verlauf der Majoritätsträgerdichte als Funktion der reziproken Temperatur für einen n-dotierten Halbleiter

Bei üblichen Betriebstemperaturen liegt die Fermienergie E_F für dotierte n-Halbleiter in dem schmalen Gebiet zwischen den Störstellenniveaus und der unteren Kante des Leitungsbands E_c. Solange $E_c - E_F > k_BT$ ist, lassen sich Berechnungen der Störstellenleitung noch in der Näherung der Boltzmannstatistik durchführen, und der Halbleiter ist nichtentartet (s. a. 8.2.2.2.). Im Falle sehr starker Dotierung mit mehr als 10^{19} Störstellen je cm^3, wie sie beispielsweise für Tunneldioden (↑ 8.3.6.2.4.) oder Halbleiterlaser (↑ 8.8.3.3.2.) verwendet wird, ist dies nicht mehr erfüllt, und es erfolgt ein Übergang zum entarteten Halbleiter, bei dem das Ferminiveau bereits innerhalb des Leitungs- oder Valenzbands liegt. Von einem *vollständig entarteten Halbleiter* spricht man, wenn der Abstand des Ferminiveaus zur Bandkante in diesen Fällen mehr als 5 k_BT beträgt.

8.3.4. Abweichungen vom thermischen Gleichgewicht

8.3.4.1. Injektion und Generation von Ladungsträgern. Bisher wurde vorausgesetzt, daß die Ladungsträgerdichten dem thermischen Gleichgewicht entsprechen, welches durch die Fermi-Dirac-Statistik und durch die Verteilung der Elektronenzustände einschließlich der Störstellenniveaus festgelegt ist. Praktisch treten jedoch häufig Abweichungen von den Gleichgewichtsdichten auf, was für die Funktion von Halbleiterbauelementen von entscheidender Bedeutung ist. Die tatsächlichen Ladungsträgerkonzentrationen n^*, p^* setzen sich dann gemäß $n^* = n + n_1$ bzw. $p^* = p + p_1$ aus den Gleichgewichtsanteilen n, p und den Beiträgen n_1, p_1 zusammen. Ursachen für Abweichungen vom thermischen Gleichgewicht bilden die *Injektion* und die *Generation* von Ladungsträgern. Unter Injektion versteht man das Einfließen zusätzlicher Ladungsträger unter der Triebkraft von elektrischen Feldern oder von Konzentrationsgradienten. Dafür reicht es schon aus, wenn der Halbleiter inhomogen dotiert ist. Zwischen den verschiedenen lokalen Bereichen bestehen dann Konzentrationsunterschiede, und es kommt zu einer Diffusion von Ladungsträgern aus Bezirken mit höherer in solche mit kleinerer Konzentration. Stärkere Inhomogenitäten sind gegeben, wenn n- und p-leitende Schichten des gleichen Halbleitermaterials oder verschiedenartiger Halbleiter aneinandergrenzen. Im Falle des gleichen Grundmaterials spricht man von pn-Übergängen (↑ 8.3.6.1.), bei Strukturen aus verschiedenartigen Halbleitern von Heteroübergängen (↑ 8.3.6.4.4.). Wegen der Unterschiede in der Verteilung und Besetzung der Elektronenzustände bauen sich an derartigen Übergängen elektrische Spannungen sowie Konzentrationsgradienten der Ladungsträger auf, wodurch ein neues Gleichgewicht hergestellt wird. Nach Anlegen äußerer Spannungen fließen dagegen Ladungsträgerströme durch den Übergang, die auch in relativ entfernten Bezirken noch erhebliche Störungen des Ladungsträgergleichgewichts bewirken können. Schließlich wird auch an Halbleiter-Metall-Kontakten häufig eine Injektion von Ladungsträgern beobachtet. Obwohl die Injektion meist einen Anstieg der Ladungsträgerkonzentrationen zur Folge hat, n_1 und p_1 also positiv sind, kann örtlich, vor allem im unmittelbaren Bereich der Übergänge, eine Verarmung von Ladungsträgern auftreten. Das führt zum Unterschreiten der Gleichgewichtsdichten mit negativen n_1, p_1 oder zur Umkehrung der Injektion in eine *Extraktion*.

Zum Unterschied von der Injektion werden bei der Generation zusätzliche Ladungsträger nicht von außen herangeführt, sondern lokal durch Energiezufuhr gebildet. Dies tritt vor allem bei der Einwirkung von elektromagnetischen oder

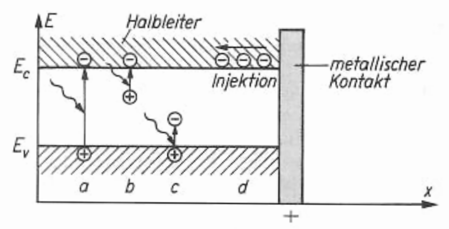


Abb. 8.3.4-1 Beispiele für die Generation von Ladungsträgern durch Lichtquanten oder Strahlungsteilchen (*a, b, c*) sowie für die Injektion aus Kontakten (*d*)

radioaktiven Strahlen auf Bezirke des Halbleiters ein. Unter Absorption von Quanten bzw. auf Kosten der kinetischen Energie der Teilchen werden entweder Ladungsträgerpaare durch zusätzliche, nichtthermische Anregung von Elektronen aus dem Valenz- in das Leitungsband geschaffen, oder es werden zusätzliche Ladungsträger aus Störstellen freigesetzt.
Abb. 8.3.4-1 veranschaulicht verschiedenartige Möglichkeiten der Injektion und Generation von Ladungsträgern in einem Halbleiter.

8.3.4.2. Rekombination von Ladungsträgern. Durch Injektion und Generation können zusätzliche Ladungsträgerkonzentrationen n_1, p_1 erzeugt werden. Wird die auslösende Ursache, z. B. eine Belichtung, abgeschaltet, so stellt sich das thermische Gleichgewicht rasch wieder ein, indem die überschüssigen Ladungsträger durch Vereinigung von Elektronen mit Löchern oder durch Rückkehr in die ionisierten Störstellen verschwinden. Dieser wichtige Vorgang bildet die Umkehrung zur Generation und wird als *Rekombination* bezeichnet. Die Rekombinationsrate, d. h. die Geschwindigkeit der Rekombination, ist meist in guter Näherung proportional zur Konzentration der zusätzlichen Ladungsträger gleicher Art, und es gilt:

$$(5)\quad \frac{dn_1}{dt} = -\frac{1}{\tau_n} n_1; \qquad \frac{dp_1}{dt} = -\frac{1}{\tau_p} p_1 .$$

Die Integration dieser Beziehungen führt auf ein exponentielles Abklingen der Überschußkonzentrationen gemäß $n_1 = n_{01} \exp[(1/\tau_n)\, n_1]$, wobei n_{01} die zusätzliche Konzentration zur Zeit $t = 0$ ist, bzw. auf eine analoge Gleichung für p_1. Aus dem Exponentialgesetz folgt die physikalische Bedeutung der in den Beziehungen 8.3.(5) eingeführten Größen τ_n, τ_p, die allerdings nur bei nicht zu großen Abweichungen vom Gleichgewicht für einen gegebenen Halbleiter Stoffkonstanten sind. Sie geben die Zeitspanne an, in der eine Überschußkonzentration n_1 oder p_1 auf den Bruchteil 1/e abfällt. Man bezeichnet τ_n bzw. τ_p als die *mittlere Lebensdauer* überschüssiger Elektronen bzw. Löcher, die Kehrwerte werden *Rekombinationskoeffizienten* genannt. Im allgemeinen sind τ_n und τ_p voneinander verschieden; die gemessenen Werte liegen im Bereich von 10^{-2} Mikrosekunden bis zu Sekunden. Auch bei gleichartigem Grundmaterial des Halbleiters kann die Lebensdauer je nach der Perfektion des Kristallgitters um Größenordnungen variieren, und zwar bewirken tiefliegende Störterme in der verbotenen Zone eine drastische Herabsetzung der Trägerlebensdauer. Solche lokalisierten Elektronenzustände werden durch Fremdatome, aber auch durch Versetzungen oder andere Kristallbaufehler geschaffen. Ferner führt der Abbruch des periodischen Kristallgitters an Oberflächen zum Auftreten spezieller *Oberflächenzustände* der Kristallelektronen, die häufig in die verbotene Zone fallen (s. a. 8.2.1.3.1.).
Nach dem von SHOCKLEY und READ aufgestellten Modell der Rekombination wirken die lokalisierten Elektronenzustände im Volumen sowie an Oberflächen als Rekombinationszentren oder Haftstellen, indem sie Überschußladungsträger einfangen und längere Zeit festhalten. Durch Übergang eines Ladungsträgers der anderen Art aus einem Band an diesen Zustand wird die Rekombination eines Elektron-Loch-Paares abgeschlossen. Dieser Mechanismus erklärt auch, warum in den Gleichungen 8.3.(5) jeweils nur die Konzentration einer Ladungsträgerart auftritt und nicht das Produkt $n_1 p_1$, wie man bei direkter Rekombination analog zur *Kinetik bimolekularer chemischer Reaktionen* (↑ 7.2.1.) erwarten sollte. Daß die direkte Vereinigung eines Elektrons mit einem Loch durch Interbandübergang unwahrscheinlicher ist als die Rekombination unter Beteiligung von Haftstellen, ist verständlich, weil die notwendige Erfüllung von Energie- und Im-

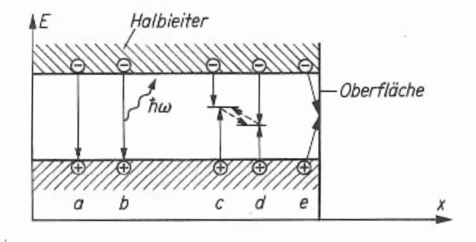


Abb. 8.3.4-2 Rekombinationsprozesse; strahlungslose und strahlende Rekombination durch Band-Band-Übergang (*a, b*), Rekombination unter Beteiligung tiefliegender Störstellen (*c, d*) sowie von Oberflächenzuständen (*e*)

pulserhaltungssatz, wie genaue Rechnungen zeigen, die Wahrscheinlichkeit für direkte Übergänge stark einschränkt.
Verschiedene Elementarprozesse der Rekombination von Ladungsträgern in dotierten Halbleitern mit Störstellen sind in Abb. 8.3.4-2 dargestellt. Wegen des häufig hohen Anteils der *Oberflächenrekombination* über Oberflächenzustände von Elektronen tritt in den oberflächennahen Gebieten vielfach eine Verarmung an Überschußladungsträgern auf, die eine Diffusion von Ladungsträgern aus dem Volumen hin zur Oberfläche bedingt. Um die Rekombination der für viele Halbleiterbauelemente wichtigen Überschußladungsträger zu minimieren, gibt es verschiedene Möglichkeiten. Zunächst sind Kristallstörungen durch Verunreinigungen und Gitterdefekte möglichst gering zu halten, was zum Einsatz von Einkristallen hoher Reinheit zwingt, wie oben schon dargelegt wurde. Daneben ist es notwendig, den Einfluß der Oberflächenrekombination zu mindern. Das geschieht einmal durch Polieren und durch Aufbringen möglichst geschlossener Deckschichten, z. B. von SiO_2 durch Oxydation von Si, die nur eine verhältnismäßig geringe Dichte von Grenzflächenzuständen zurücklassen. Ferner sorgt man durch geeignete konstruktive Gestaltung von Transistoren und Schaltkreisen dafür, daß die relative Oberfläche der kritischen Halbleiterbereiche, insbesondere der Basiszonen von Transistoren, möglichst klein ist.

8.3.4.3. Strahlende und strahlungslose Übergänge. Man fragt schließlich nach dem Verbleib der Energie, die bei der Rekombination von Ladungsträgern freigesetzt wird. In Umkehrung der Generation von Ladungsträgern durch Absorption von Lichtquanten, also Elektron-Photon-Wechselwirkung, könnte man erwarten, daß diese Energie in Form von Quanten elektromagnetischer Strahlung freigesetzt wird. Tatsächlich wird eine Lichtemission im Zusammenhang mit Rekombinationsprozessen beobachtet, worauf die Funktion von Fotodioden und Halbleiterlasern beruht (s. a. 8.8.3.). Zu beachten ist dabei, daß die Energiedifferenzen bei üblichen Halbleitern nur zur Bildung von Quanten im Infraroten ausreichen. Sichtbares Licht kann nur emittiert werden, wenn Interbandübergänge eine wesentliche Rolle spielen und der Bandabstand wenigstens 2,1 eV beträgt. *Strahlende Übergänge* mit Emission im sichtbaren Spektralgebiet lassen sich daher nur bei Einsatz von Halbleitern mit großem Bandabstand, z. B. GaP, realisieren. Häufig überwiegen jedoch sowohl im infraroten als auch im sichtbaren Bereich *strahlungslose Übergänge*, oder es wird sogar keinerlei Lichtemission festgestellt. Erst in neuerer Zeit konnte gezeigt werden, daß für das Ausbleiben der Emission von elektromagnetischen Strahlen Multiphononenprozesse verantwortlich sind. Durch Elektron-Phonon-Wechselwirkung (s. a. 8.2.3.6.1.) wird die freiwerdende Energie dabei an das Gitter übertragen, also letztlich in Wärme überführt, indem viele Phononen je Rekombinationsvorgang erzeugt werden.

8.3.4.4. Ladungsträgerbilanz und spezielle Fälle. Bisher wurde die Rekombination unter der Voraussetzung untersucht, daß die Ursachen für Abweichungen vom Ladungsträgergleichgewicht nicht mehr wirken. Im allgemeinen finden Injektion, Generation und Rekombination jedoch gleichzeitig statt. Die Ladungsträgerkonzentrationen, die sich dann einstellen, ergeben sich aus der Lösung von *Bilanzgleichungen*. Diese besitzen die Form $\partial n^*/\partial t = i_n + g_n - r_n$ bzw. $\partial p^*/\partial t = i_p + g_p - r_p$. Die linken Seiten der Gleichungen geben die Rate der zeitlichen Änderung der gesamten Trägerdichten $n^* = n + n_1$ bzw. $p^* = p + p_1$ an. Auf der rechten Seite stehen die Injektionsraten i_n, i_p, die Generationsraten g_n, g_p und die Rekombinationsraten r_n, r_p, die selbst Funktionen der Trägerdichten und des Orts sein können. Die Bilanzgleichungen sagen einfach aus, daß die zeitliche Änderung der Ladungsträgerkonzentrationen an einem bestimmten Ort des Halbleiters gleich der Differenz zwischen hinzukommenden und verschwindenden Ladungsträgern sein muß. Im übrigen sind die Gleichungen für die beiden Trägerarten noch über die Forderung nach *Quasineutralität* in nicht zu kleinen Bezirken des Halbleiters miteinander gekoppelt (s. a. 8.3.3.4.). Ohne auf Einzelheiten der Rechnung einzugehen, sollen nachfolgend einige Ergebnisse für wichtige Spezialfälle aufgeführt werden:

8.3.4.4.1. *Stationärer Fall.* Bleiben Injektions- und Generationsrate über längere Zeit konstant, so stellt sich nach einem Einlaufvorgang ein stationärer Zustand ein. Die Rekombinationsrate wird dann gleichfalls konstant, und die zeitlichen Differentialquotienten auf den linken Seiten der Bilanzgleichungen verschwinden. Trotzdem sind die Ladungsträgerkonzentrationen von denen des Gleichgewichtsfalls verschieden; gewöhnlich liegen sie darüber, und es gilt $n^* > n$, $p^* > p$ bzw. $n_1 > 0$, $p_1 > 0$. Nach Aufhören der Injektion und Generation klingen die zusätzlichen Ladungsträgerdichten n_1, p_1 nach den Beziehungen 8.3.(5) exponentiell ab. Die Geschwindigkeit, mit der die Annäherung an den Gleichgewichtsfall erfolgt, hängt von den mittleren Trägerlebensdauern τ_n, τ_p ab (↑ 8.3.4.2.).

8.3.4.4.2. *Zeitliches Abklingen einer Raumladung.* Im Nichtgleichgewichtsfall kann sich in einem Halbleiter örtlich eine Raumladung ausbilden. Die Forderung nach Neutralität wird dann in eine nach Quasineutralität abgemildert (↑ 8.3.3.4.). Allerdings bleibt der Überschuß von

Ladungen eines Vorzeichens fast stets klein gegen die gesamte Ladung je Volumeneinheit. Wird nun die Ursache für die Abweichung von der örtlichen Neutralität abgestellt, so klingt die Raumladung mit der Dichte ϱ_0 entsprechend $\varrho = \varrho_0 \exp -[t/\tau]$ ab. Die hier auftretende dielektrische oder Maxwellsche Relaxationszeit τ, die nicht mit den Trägerlebensdauern verwechselt werden darf, liegt bei Halbleitern in den Größenordnungen von 10^{-6} bis 10^{-18} s. Raumladungen werden also sehr rasch abgebaut.

8.3.4.4.3. *Debyelänge.* Wird in einem eindimensionalen Halbleiter an der Stelle $x = 0$ eine Raumladung ϱ_0 stationär aufrechterhalten, so klingt diese Abweichung von der Neutralität mit wachsender Entfernung x gleichfalls exponentiell ab (Abb. 8.3.4-3). Es gilt $\varrho(x) = \varrho_0 \exp -[x/L_D]$. Die hier eingeführte Debyelänge

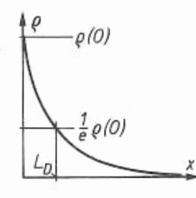


Abb. 8.3.4-3 Abklingen einer Raumladung bei $x = 0$ mit wachsendem Abstand x von der Grenzfläche; L_D Debyelänge

L_D gibt an, in welchem Abstand die örtliche Raumladung auf $1/e$ abnimmt. Beispielsweise beträgt L_D für mäßig dotiertes Silicium bei Zimmertemperatur etwa 50 nm. Raumladungen treten also örtlich nur in begrenzten Bereichen, hauptsächlich in der Nähe von Kontakten, *pn*-Übergängen oder Oberflächen auf.

8.3.5. Stromfluß in Halbleitern

8.3.5.1. Stromanteile im Halbleiter. Wird an einen Halbleiter über metallische Kontakte eine elektrische Spannung angelegt, beginnen die Ladungsträger unter dem Einfluß des örtlich aufgebauten elektrischen Felds $\boldsymbol{E}$ zu wandern. Die hierdurch hervorgerufenen *Driftströme* zerlegt man zweckmäßigerweise in den Teilstrom, der von den Elektronen getragen wird, und in den Teilstrom der Löcher, welcher in entgegengesetzter Richtung fließt. Mit Einführung der *Stromdichten* j_n' für den Anteil der Elektronen im Leitungsband und j_p' für den Anteil der Löcher im Valenzband gilt also für den gesamten Driftstrom $j_{\text{ges}}' = j_n' + j_p'$. Strenggenommen sind die Stromdichten Vektoren; wir beschränken uns jedoch auf den eindimensionalen Fall und können sie dann als Skalare auffassen, deren Betrag die durch eine Flächeneinheit fließende elektrische Stromstärke angibt. Dabei liefern die Elektronen und Löcher trotz entgegengesetzter Bewegungsrichtung wegen der gleichfalls entgegengesetzten Ladung einen positiven Beitrag.

Außer bei sehr hohen Feldstärken wird in Halbleitern das Ohmsche Gesetz befolgt, wonach die Stromdichte proportional zur Feldstärke ist, d. h., es gilt $j_{\text{ges}}' = \sigma_{\text{ges}}|E| = (\sigma_n + \sigma_p)\,|\boldsymbol{E}|$. Die elektrische Leitfähigkeit, die hier gleichfalls in zwei Anteile zerlegt wurde, läßt sich analog zu Gleichung 8.2.(12) durch die Ladungsträgerbeweglichkeiten μ_n und μ_p ausdrücken (↑ 8.2.3.6.). Für die Anteile des Driftstroms folgt dann $j_n' = en^*\mu_n|E|$ und $j_p' = ep^*\mu_p|E|$ und die gesamte Leitfähigkeit ist $\sigma_{\text{ges}} = e(n^*\mu_n + p^*\mu_p)$. Die Leitfähigkeit wird somit von den Ladungsträgerkonzentrationen n^*, p^* und den Beweglichkeiten μ_n, μ_p bestimmt. Im Fall des thermischen Gleichgewichts, der bei nicht zu großen Feldstärken und homogenen Halbleitern erfüllt ist, können die Ladungsträgerdichten n^*, p^* durch die Gleichgewichtswerte n, p ersetzt werden.

Andererseits sind bei inhomogenen Halbleiterstrukturen, wie sie in elektronischen Bauelementen gewöhnlich vorliegen, zwei weitere Stromanteile zu berücksichtigen, nämlich die *Diffusionsstromdichten* j_n'' und j_p'' der Elektronen bzw. der Löcher. Diffusionsströme treten auf, wenn die Ladungsträgerkonzentrationen örtlich verschieden sind. Die Ladungsträger sind dann ähnlich wie Gasteilchen bestrebt, die Konzentrationsunterschiede auszugleichen, in dem sie von Stellen höherer zu solchen niederer Trägerdichte diffundieren. Nach dem *1. Fickschen Gesetz der Diffusion* (↑ 6.2.5.3.) sind die Teilchenstromdichten proportional zu den Konzentrationsgradienten, die in einer Dimension als Differentialquotienten $\mathrm{d}n^*/\mathrm{d}x$ bzw. $\mathrm{d}p^*/\mathrm{d}x$ anzusetzen sind. In drei Dimensionen sind entsprechend die räumlichen Gradienten $\operatorname{grad} n = \mathrm{d}n^*/\mathrm{d}\boldsymbol{r}$ bzw. $\operatorname{grad} p = \mathrm{d}p^*/\mathrm{d}\boldsymbol{r}$ zu benutzen. Nach Multiplikation mit der Elementarladung $-e$ bzw. $+e$ erhält man für die Diffusionsstromdichten der Elektronen bzw. Löcher $j_n'' = eD_n\,\mathrm{d}n^*/\mathrm{d}x$ bzw. $j_p'' = -eD_p\,\mathrm{d}p^*/\mathrm{d}x$. Die Stoffkonstanten D_n, D_p werden als *Diffusionskoeffizienten* der Elektronen bzw. Löcher bezeichnet. Hinsichtlich der Vorzeichengebung ist zu beachten, daß beide Trägerarten in Richtung abnehmender Konzentration diffundieren, die Gradienten aber gerade in umgekehrter Richtung positiv gerechnet werden. Da bei gleicher Bewegungsrichtung ferner Ladungen entgegengesetzten Vorzeichens transportiert werden, tragen die Diffusionsstromdichten verschiedene Vorzeichen.

Insgesamt setzt sich die Stromdichte j(total) in Halbleitern demnach aus vier Anteilen zusammen, die sich zu den Teilstromdichten j_n(total) der Elektronen im Leitungsband und j_p(total) der Löcher im Valenzband vereinigen lassen. Unter Weglassen der Klammerbezeichnungen

ergibt sich das folgende *System allgemeiner Trägertransportgleichungen*:

(6) $j = j_n + j_p$;
$j_n = en^*\mu_n |E| + eD_n(dn^*/dx)$
$j_p = ep^*\mu_p |E| - eD_p(dp^*/dx)$.

Zusammen mit den *Bilanzgleichungen* (↑ 8.3.4.4.) und den grundlegenden *Maxwellgleichungen* der Elektrodynamik (s. a. 2.2.4.) gestatten es diese Gleichungen im Prinzip, das elektrische Verhalten elektronischer Halbleiterbauelemente vollständig zu beschreiben. Praktisch kann die Berechnung zu erheblichen Schwierigkeiten führen, weil die Coulombkräfte zwischen den geladenen Störstellen, Elektronen und Löchern eine Kopplung zwischen den einzelnen Stromanteilen bewirken. Beispielsweise führt die Forderung nach Quasineutralität dazu, daß sich lokale Anhäufungen von Ladungsträgern, wie sie etwa nach Belichten durch eine Spaltblende auftreten, nicht durch unabhängige Diffusionsströme, sondern in Form einer *ambipolaren Diffusion* ausbreiten. Die schnelleren Ladungsträger schleppen die mit der kleineren Diffusionskonstante gewissermaßen hinter sich her, und es stellt sich ein Transport mit einem *ambipolaren Diffusionskoeffizienten* D_a ein, der zwischen D_n und D_p liegt. Bei deutlichem Überwiegen einer Ladungsträgerart, also bei stärker dotierten Halbleitern, kann D_a näherungsweise gleich dem Diffusionskoeffizienten der Minoritätsträger gesetzt werden.

Außerdem ist eine genaue Kenntnis der elektrischen Stoffkonstanten von Halbleitern notwendig, die im allgemeinen experimentell bestimmt werden müssen. Einige Zahlenwerte für wichtige Halbleiter sind in Tab. 8.3.1-2 zusammengestellt. Zwischen den Diffusionskoeffizienten und den Beweglichkeiten besteht allerdings der eindeutige Zusammenhang

(7) $D_n = (k_BT/e)\mu_n$ bzw. $D_p = (k_BT/e)\mu_p$

(k_B Boltzmannkonstante, T absolute Temperatur, e Elementarladung), der nach dem Entdecker *Einsteinbeziehung* genannt wird. Man braucht daher nur eine der Größen zu bestimmen und wählt hierfür gewöhnlich die leichter zu messende Beweglichkeit.

8.3.5.2. Messung elektrischer Parameter von Halbleitern. Obwohl ein näheres Eingehen auf die physikalische Meßtechnik im Rahmen dieses Buchs nicht möglich ist, sollen an dieser Stelle einige Bemerkungen zur Messung der Ladungsträgerkonzentrationen und der Beweglichkeiten eingefügt werden. Relativ am einfachsten liegen die Verhältnisse bei homogen und ausreichend stark dotierten Halbleiterproben, bei denen Diffusionsströme entfallen und der Stromtransport durch eine Ladungsträgerart bestimmt wird. Als Beispiel sei ein n-Halbleiter betrachtet, in dem der Stromfluß fast ausschließlich durch den Driftanteil der Elektronen als Majoritätsträger erfolgt. Da die Löcher als Minoritätsträger schon bei mäßig starker Dotierung nach dem *Massenwirkungsgesetz* $np =$ const. (↑ 8.3.2.2.) in einer um Größenordnungen geringeren Konzentration vorhanden sind, kann ihr Beitrag vernachlässigt werden. Nach den Beziehungen 8.3.(6) folgt für diesen Fall $j \cong j_n = en\mu_n |E| = \sigma |E|$ oder $\varkappa = 1/\sigma = 1/(en\mu_n)$. Aus einer Messung des spezifischen elektrischen Widerstands $\varkappa$ läßt sich daher zunächst nur das Produkt aus Ladungsträgerkonzentration und Beweglichkeit ermitteln. Um diese Größen getrennt zu erhalten, braucht man eine weitere Information. Diese läßt sich relativ einfach und genau durch Messung des *Halleffekts* gewinnen.

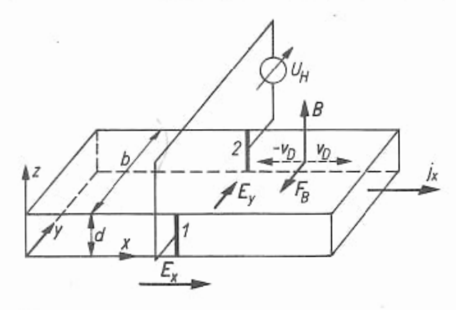


Abb. 8.3.5-1 Deutung und Messung des Halleffekts (Erläuterung im Text)

Das Prinzip des nach HALL benannten Effekts ist in Abb. 8.3.5-1 dargestellt. Eine stromdurchflossene Probe wird einem starken Magnetfeld $\boldsymbol{B}$ ausgesetzt, dessen Feldlinien senkrecht zur Stromrichtung stehen. Die Ladungsträger unterliegen senkrecht zu ihrer mittleren Driftgeschwindigkeit $\boldsymbol{v}_D$ und zur magnetischen Induktion $\boldsymbol{B}$ der ablenkenden Lorentzkraft, die durch $\boldsymbol{F}_B = \pm e(\boldsymbol{v}_D \times \boldsymbol{B})$ gegeben ist (s. a. 2.2.3.). Hier steht das Pluszeichen für Löcher, das Minuszeichen für Elektronen. Da sich jedoch auch $\boldsymbol{v}_D$ für die beiden Ladungsträgerarten im Vorzeichen unterscheidet, werden Elektronen und Löcher nach der gleichen Richtung abgelenkt. Gerade dadurch laden sich die Seitenflächen der Probe für Elektronen bzw. Löcher entgegengesetzt auf. Hierdurch baut sich bei offenen Kontakten *1* und *2* ein Gegenfeld auf, das der weiteren Ablenkung von Ladungsträgern ein Ende setzt. Im Kräftegleichgewicht gilt $\boldsymbol{F}_B \pm e\boldsymbol{E}_H = 0$, wenn $\boldsymbol{E}_H$ die Hallfeldstärke quer zur Stromrichtung bezeichnet und Plus- bzw. Minuszeichen wieder für Löcher bzw. Elektronen stehen. E_H wird mit Hilfe eines hochohmigen Spannungsmeßgeräts über die Kontakte *1* und *2* bestimmt. Für die Hallspannung U_H läßt sich auch $U_H = E_Hb = R_HIB/d$ schreiben, wobei b

die Breite und d Dicke der Probe, I die Stromstärke, B die magnetische Induktion und R_H den *Hallkoeffizienten* bedeuten. Das Vorzeichen von R_H ist bei vorherrschender n-Leitung negativ, bei p-Leitung positiv. Allein aus dem Vorzeichen kann also bereits auf den Leitungscharakter geschlossen werden. Die Lösung der eingangs gestellten Aufgabe ist mit Hilfe von Näherungsausdrücken für R_H möglich. Im Falle starker Magnetfelder gilt bei p-Leitung $R_{Hp} = +1/(ep)$, bei n-Leitung $R_{Hn} = -1/(en)$. Demzufolge lassen sich aus den Hallkoeffizienten die Trägerkonzentrationen und zusammen mit den Widerstandsmeßdaten die Beweglichkeiten bestimmen. Bei schwächeren Magnetfeldern ist freilich noch ein Korrekturfaktor zu berücksichtigen. Falls beide Ladungsträgerarten zum Stromfluß beitragen, der Halbleiter also nicht allzuweit vom Zustand der Eigenleitung entfernt ist, gelten für den Hallkoeffizienten erheblich kompliziertere Ausdrücke. Trotzdem ist es vielfach möglich, die Beweglichkeiten und Trägerkonzentrationen aus Halleffekt- und Widerstandsmessungen zu bestimmen, indem man verschieden dotierte Proben vergleicht.

8.3.5.3. Beweglichkeit und Streuung der Ladungsträger. Die Gültigkeit des Ohmschen Gesetzes verlangt, daß die Beweglichkeiten und die Trägerdichten nicht von der Feldstärke abhängen. Die Driftbewegung der Ladungsträger erfolgt mit den mittleren Driftgeschwindigkeiten v_{Dn} bzw. v_{Dp}, die proportional zur Feldstärke sind (s. a. 8.2.3.6.). Das ist ein wichtiger Unterschied zum Verhalten freier Elektronen im Vakuum, z. B. in einer Fernsehbildröhre, deren Geschwindigkeit zunimmt, solange ein elektrisches Feld einwirkt. Tatsächlich werden auch die Ladungsträger im Halbleiter durch ein elektrisches Feld ständig beschleunigt, jedoch geht die zusätzlich zur thermischen Bewegung erlangte Driftgeschwindigkeit im Gefolge von Streuprozessen nach statistisch verteilten Wegstrecken wieder verloren. Die im Feld erlangte Zusatzgeschwindigkeit weist in Abhängigkeit von der Zeit den in Abb. 8.3.5-2 dargestellten Verlauf auf. Der zeitliche Mittelwert ist die Driftgeschwindigkeit des Ladungsträgers, und die gemittelte Zeit zwischen zwei Stoßprozessen wird *mittlere freie Flugzeit* oder *Relaxationszeit* τ_n bzw. τ_p genannt. Als Produkte $l_n = v_{Dn}\tau_n$ bzw. $l_p = v_{Dp}\tau_p$ erhält man schließlich die *mittleren freien Weglängen* der Elektronen bzw. Löcher. Diese Größen hängen also unmittelbar

Abb. 8.3.5-2 Momentane Zusatzgeschwindigkeit v_Z eines Kristallelektrons im elektrischen Feld; mittlere Driftgeschwindigkeit v_D *gestrichelt*

mit den Beweglichkeiten zusammen, die ja als Verhältnisse der Driftgeschwindigkeiten zur Feldstärke gemäß $\mu_n = v_{Dn}/|E|$ bzw. $\mu_p = v_{Dp}/|E|$ definiert sind (↑ 8.2.3.6.). Aus Tab. 8.3.1-2 entnimmt man, daß die Beweglichkeiten für verschiedene Halbleiter sehr unterschiedlich sind, wobei für Elektronen meist größere Werte gefunden werden als für Löcher. Die angegebenen Werte beziehen sich auf Zimmertemperatur und weitgehend perfekte, schwach dotierte Einkristalle.

Der Temperaturgang der Beweglichkeiten wird vom dominierenden Streumechanismus der Ladungsträger im Gitter bestimmt. Bei Halbleitern schwacher Dotierung und hoher Kristallperfektion überwiegt die Elektron-Phonon-Steuung (s. a. 8.2.3.6.), deren Wahrscheinlichkeit infolge zunehmender Anfachung der Gitterschwingungen mit der Temperatur wächst. Infolgedessen sinkt die von diesem Mechanismus bestimmte *Gitterbeweglichkeit* mit zunehmender Temperatur proportional T^{-z}, wobei $z = 3/2 \dots 5/2$ ist. Im Falle stärkerer Dotierung werden die Ladungsträger wegen der elektrostatischen Wechselwirkung vor allem auch an den ionisierten Störstellen gestreut. Diese *Störstellenbeweglichkeit* steigt mit der Temperatur etwa wie $T^{3/2}$ an, denn bei kleinen T verweilen die Ladungsträger wegen der geringeren thermischen Geschwindigkeit länger in der Nähe der Streuzentren, und sie werden stärker abgelenkt. Weiterhin kann die Beweglichkeit der Ladungsträger im Halbleiter durch die Beeinflussung der Elektronen untereinander, durch Streuung an neutralen Störstellen, an Versetzungen und anderen Gitterdefekten eingeschränkt werden. Die letzten Effekte treten naturgemäß um so stärker hervor, je geringer die Kristallperfektion ist; das gilt beispielsweise für Halbleiterschichten mit vielen Versetzungen, die als *heteroepitaktische Schichten* auf isolierenden Substraten anderer Stoffe wie Saphir oder Spinell abgeschieden werden. Bei *dünnen Schichten* ist daneben generell eine geringere Leitfähigkeit zu beobachten, weil durch Streuung an der Oberfläche sowie der Grenzfläche Schicht/Substrat die mittlere freie Weglänge herabgesetzt wird.

Beim Zusammenwirken mehrerer Streumechanismen müßte man eine reziproke freie Flugzeit der Ladungsträger $1/\tau$ ermitteln, die annähernd der Summe der reziproken Anteile $\sum_i (1/\tau_i)$ entspricht, die von den einzelnen Mechanismen bewirkt werden. Falls ein Streumechanismus überwiegt, kann man daraus auch die häufig benutzte Näherung $1/\mu = \sum_i (1/\mu_i)$ begründen, nach der sich die Gesamtbeweglichkeit als reziproke Summe aus den Anteilen zusammensetzt. Aus der gegenläufigen Temperaturab-

hängigkeit der Gitter- und der Störstellenbeweglichkeit folgt, daß μ_n sowie μ_p mit der Temperatur häufig ein Maximum durchlaufen (Abb. 8.3.5-3). Für den praktisch wichtigsten Temperaturbereich von 0 bis 100 °C weisen die Ladungsträgerbeweglichkeiten in stärker dotiertem Silicium nur einen geringen Gang auf.

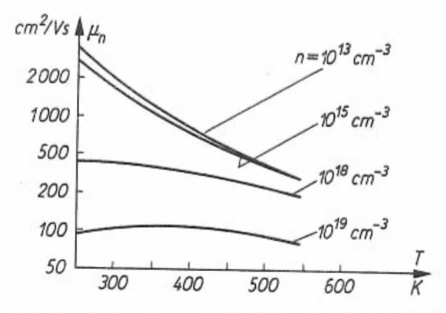


Abb. 8.3.5-3 Beweglichkeit der Elektronen μ_n für verschieden stark dotiertes *n*-Silicium als Funktion der Temperatur *T*; *n* Majoritätsträgerdichte

Für elektronische Bauelemente sind hohe Beweglichkeiten erwünscht, weil die endliche Laufzeit der Ladungsträger die Reaktion bei zeitlichen Änderungen der elektrischen Parameter, also das Verhalten bei sehr hohen Frequenzen, begrenzt. Neuerdings versucht man daher, die höhere Beweglichkeit von gekühlten und nur schwach dotierten Halbleiterschichten aus Materialien wie GaAs für sehr schnell schaltende *MOS-Kurzkanal-Transistoren* (s. a. 8.3.6.4.) zu nutzen. Die Versorgung der

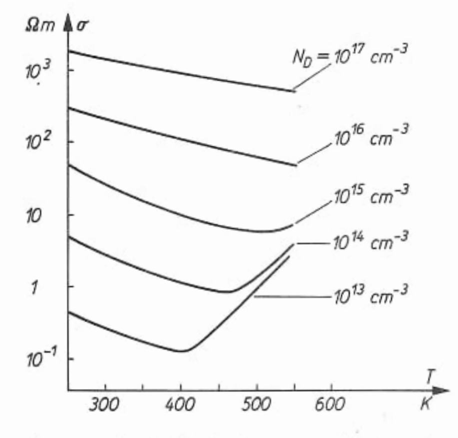


Abb. 8.3.5-4 Leitfähigkeit von n-Silicium als Funktion von *T* im mittleren Temperaturbereich für verschiedene Donatorkonzentrationen N_D

Schichten mit Ladungsträgern geschieht dabei durch Eindiffusion aus einer benachbarten, stärker dotierten Schicht.

8.3.5.4. Leitfähigkeit als Funktion der Temperatur. Der Temperaturverlauf der Leitfähigkeit $\sigma(T)$ wird durch Überlagerung der Temperaturabhängigkeiten der Ladungsträgerkonzentrationen und der Beweglichkeiten bestimmt. Bei tiefen Temperaturen steigt σ mit wachsendem T bei dotierten Halbleitern zunächst steil an, weil die exponentielle Zunahme der Ladungsträgerdichte den umgekehrten Gang der Gitterbeweglichkeit überwiegt. Danach bleibt die Ladungsträgerkonzentration infolge Störstellenerschöpfung etwa gleich, so daß der Verlauf der Leitfähigkeit durch den der Beweglichkeit bestimmt wird. Meist wird eine leicht fallende Leitfähigkeit beobachtet, da bei mittleren Temperaturen die Streuung der Ladungsträger an Phononen den Einfluß der Störstellenbeweglichkeit übertrifft (Abb. 8.3.5-4). Bei noch höheren Temperaturen verursacht die thermische Anregung von Elektronen aus dem Leitungsband schließlich einen steilen Anstieg der Leitfähigkeit, die schließlich auch bei dotierten Halbleitern durch die Eigenleitung bestimmt wird.

8.3.5.5. Abweichungen vom Ohmschen Gesetz bei hohen Feldstärken. Bei hohen elektrischen Feldstärken ($|E| > 10^6\ \mathrm{Vm^{-1}}$) werden in zunehmendem Maße Abweichungen vom Ohmschen Gesetz beobachtet. Zunächst steigt die Driftgeschwindigkeit der Ladungsträger nicht mehr proportional zu $|E|$, sondern langsamer an, womit die Beweglichkeiten als Proportionalitätskonstanten ihren Sinn verlieren. Darüberhinausgehend erlangen die Kristallelektronen bei extrem hohen Feldstärken von $10^7\ \mathrm{Vm^{-1}}$ zwischen zwei Streuprozessen so viel Bewegungsenergie, daß sie selbst imstande sind, weitere Elektron-Loch-Paare durch Interbandübergänge zu bilden. Dadurch kommt es zu einer rasch anschwellenden Ladungsträgervervielfachung, die zunächst zu reversiblen *Lawinendurchbrüchen* und schließlich zu einem meist von irreversiblen strukturellen Veränderungen begleiteten elektrischen Durchschlag führt. Zeitliche Instabilitäten im Gebiet der reversiblen Durchbrüche werden in den *Lawinen-Laufzeit-Dioden* oder *IMPATT-Dioden* zur Erzeugung elektrischer Schwingungen im Bereich höchster Frequenzen von einigen 10^{12} Hz genutzt.

Noch vor Einsetzen der Lawinendurchbrüche können zwei andere Effekte auftreten, die ebenfalls für spezielle Bauelemente wichtig sind. Der *Zenereffekt* beruht auf einer so starken Bandverbiegung an Grenzflächen oder *pn*-Übergängen, daß Elektronen aus dem Valenzband in das Leitungsband gelangen können, indem sie die verbotene Zone durchtunneln.

Dies ist nach dem *quantenmechanischen Tunneleffekt* (↑ 2.5.4.8.) möglich, wenn die unbesetzten Zustände des Leitungsbands in genügend kleinen Abständen energetisch tiefer liegen als die oberen Zustände des Valenzbands. Darauf beruhende Zenerdioden dienen vor allem zur Spannungsbegrenzung. Beim *Gunneffekt*, der an Halbleitern mit mehreren Minima im $E(k)$-Diagramm des Leitungsbands auftritt (s. a. 8.2.1.4.3.), beispielsweise bei GaAs, werden Elektronen nach ausreichender Energieaufnahme im Feld in energetisch höher liegende Nebenminima angeregt (Abb. 8.3.5-5). Dort finden sie jedoch Bedingungen vor, die auf eine wesentlich kleinere Beweglichkeit als in den Ausgangszuständen hinauslaufen. Dies bedingt eine Art „Stau", der zu Strominstabilitäten Anlaß gibt, die in Gunndioden für Oszillatoren im Höchstfrequenzbereich ausgenutzt werden können.

8.3.6. Physikalische Grundlagen der Halbleiterbauelemente

8.3.6.1. *pn*-Übergang. Durch geeignete Dotierung mit Donator- und Akzeptorstörstellen (s. a. 8.3.3.1.) lassen sich in einem Halbleiterkristall *p*- und *n*-leitende Bereiche schaffen, die durch eine schmale Übergangszone von weniger als 1 μm Dicke getrennt sind. Derartige *pn*-Übergänge bilden die elementare Funktionseinheit der bipolaren Halbleiter- und Mikroelektronik; sie sind aber auch für unipolare MOS-Bauelemente und -Schaltkreise bedeutungsvoll. Die Kennzeichnung *Bipolartechnik*

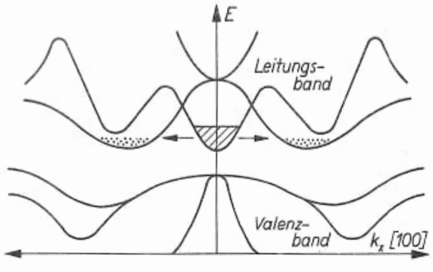


Abb. 8.3.5-5 Bandzweige von Galliumarsenid in [100]-Richtung und Übergänge (*Pfeile*), die zum Gunneffekt führen

Abb. 8.3.6-1 *pn*-Übergang in Silicium; *links* im stromlosen Zustand, *Mitte* bei Vorspannung in Sperrichtung; a) Verlauf der Störstellendichten in cm^{-3}, b) Ladungsträgerdichten in cm^{-3}, c) Verlauf des Potentials φ in V, d) Stromdichten in $mA\ cm^{-2}$

soll zum Ausdruck bringen, daß die Funktionsweise durch den Beitrag beider Ladungsträgerarten, Elektronen und Löcher, geprägt wird. Als Grundlage für das Verständnis der elektronischen Halbleiterbauelemente wird im folgenden das Verhalten von *pn*-Übergängen im stromlosen und im stromführenden Zustand diskutiert.

8.3.6.1.1. *Stromloser Zustand des pn-Übergangs.* Der stromlose Fall liegt vor, wenn an den *pn*-Übergang, der aus homogen dotierten Bereichen bestehen soll, keine äußere Spannung angelegt wird. In diesem Fall stellt sich ein Gleichgewichtszustand ein, indem am Übergang ein elektrisches Potentialgefälle $\Delta\varphi = U_0$ aufgebaut wird. U_0 nennt man die *Leerlauf-* oder *Diffusionsspannung* des *pn*-Übergangs. Die Potentialdifferenz am Übergang führt dazu, daß sich die Fermienergie E_F, die bei getrennten *p*- und *n*-Bereichen unterschiedliche Werte hätte, auf beiden Seiten des *pn*-Übergangs auf ein einheitliches Niveau einstellt. Die Konstanz des Ferminiveaus ist ganz allgemein das Kennzeichen für den Gleichgewichtszustand.

Das Zustandekommen der Potentialdifferenz U_0 läßt sich durch eine dynamische Betrachtung verständlich machen. Die unterschiedliche Dotierung mit einer Konzentration von N_A Akzeptoren bzw. N_D Donatoren führt zu einem Gradienten der Ladungsträgerdichten n und p am Übergang (Abb. 8.3.6-1). Infolgedessen diffundieren Elektronen von der *n*-Seite, wo sie Majoritätsträger sind, nach der *p*-Seite und umgekehrt Löcher von der *p*- zur *n*-Seite. Zwischen diesen durchdiffundierten Ladungsträgern und den zurückgebliebenen ionisierten Störstellen baut sich ein elektrisches Feld $\boldsymbol{E}$ auf, das seinerseits Driftströme in entgegengesetzter Richtung hervorruft. Gleichgewicht ist erreicht, wenn sich Drift- und Diffusionsströme gerade kompensieren. Die Teilstromdichten j_n und j_p müssen dann jeder für sich Null werden, und nach den Gleichungen 8.3.(6) folgt $[e/(k_B T)]\, E = -(\mathrm{d}n/\mathrm{d}x)\, n$ bzw. $[e/(k_B T)]\, E = (\mathrm{d}p/\mathrm{d}x)\, p$, wobei von der Einsteinbeziehung $\mu = e/(k_B T)\, D$ Gebrauch gemacht wurde. Drückt man die elektrische Feldstärke definitionsgemäß als negativen Gradienten des elektrischen Potentials $E = -\mathrm{d}\varphi/\mathrm{d}x$ aus, so lassen sich die vorstehenden Gleichungen über die Gesamtbreite der Raumladungszone integrieren, und man erhält für die Elektronen $[e/(k_B T)]\,(\varphi_n - \varphi_p) = [e/(k_B T)]\, U_0 = \ln\,(n_n/n_p)$ sowie eine analoge Gleichung für die Löcher. Dabei weisen die Indizes n und p auf die beiden Seiten des *pn*-Übergangs hin, und die Werte des Potentials sowie der Trägerdichten entsprechen den konstanten Bereichen außerhalb der Raumladungszone (↑ Abb. 8.3.6-1). Durch beidseitiges Erheben zur Exponentialfunktion ergibt sich schließlich

(8) $n_p = n_n \exp - [eU_0/(k_B T)]$ bzw.
$p_n = p_p \exp\,[eU_0/(k_B T)]$,

womit eine eindeutige Beziehung zwischen der Leerlaufspannung und den Ladungsträgerdichten zu beiden Seiten des *pn*-Übergangs hergestellt ist.

8.3.6.1.2. *pn-Übergang bei angelegter Spannung.* Wird an den *pn*-Übergang über geeignete Kontakte eine Spannung angelegt und die Verbindung mit einem äußeren Stromkreis hergestellt, so kann Stromfluß eintreten, der aus Abweichungen vom Gleichgewichtszustand resultiert. Je nach der Polarität der angelegten Spannung sind dabei zwei verschiedene Fälle zu unterscheiden.

Durchlaßrichtung (↑ Abb. 8.3.6-1) liegt vor, wenn der positive Pol an der *p*-Seite und der negative Pol an der *n*-Seite angelegt werden, was man sich leicht merken kann. Die zusätzliche Spannung U am Übergang erniedrigt die Potentialdifferenz U_0 des stromlosen Zustands, so daß die Majoritätsträger den Übergang leichter passieren können. Die Diffusionsströme überwiegen dann über die rücktreibenden Driftströme. Da die Durchtrittswahrscheinlichkeit exponentiell von der Potentialdifferenz abhängt, verursachen schon relativ kleine Vorspannungen in der Größenordnung von 1 V einen beachtlichen Stromfluß. Hierdurch wird von beiden Seiten aus eine Injektion von Ladungsträgern bewirkt, die zu einem Anstieg der Minoritätsträgerdichten über die Gleichgewichtswerte n_p bzw. p_n führt. Demgegenüber macht sich der Abfall von n_n bzw. p_p wegen der um Größenordnungen höheren Majoritätsträgerkonzentrationen kaum bemerkbar.

Unter der Triebkraft des Konzentrationsgefälles sowie des elektrischen Felds wandern die injizierten Ladungsträger vom *pn*-Übergang weg. Dabei finden Rekombinationsprozesse (↑ 8.3.4.2.) statt, wodurch die Dichten der überschüssigen Ladungsträger mit zunehmender Entfernung vom Übergang abnehmen, bis im Falle genügend ausgedehnter Halbleiter wieder die Gleichgewichtskonzentrationen n_p bzw. p_n erreicht werden. Indessen muß die Gesamtstromdichte $j = j_n + j_p$ aus Gründen der Kontinuität in jedem Querschnitt des stromführenden Halbleiters konstant sein. In der Nähe des Übergangs wird er durch die Stromanteile der durchtretenden Träger bestimmt. Mit zunehmendem Abstand vom Übergang nehmen die Strombeiträge der injizierten Träger ab, und der Strom wird schließlich in erster Linie von den Majoritätsträgern getragen, die vom Feld in Richtung des Übergangs getrieben werden.

Bei einem in *Sperrichtung* vorgespannten *pn*-Übergang, d. h. bei Anlegen des positiven Pols an der *n*-Seite und umgekehrt, wird die Potentialbarriere am Übergang gegenüber dem

stromlosen Zustand erhöht und der Durchtritt von Majoritätsträgern praktisch unterbunden (↑ Abb. 8.3.6-1). Daß trotzdem durch einen gesperrten *pn*-Übergang ein kleiner Sperr- oder Reststrom fließt, ist den Minoritätsträgern zuzuschreiben, die bei dieser Polung wegen ihrer entgegengesetzten Ladung günstige Durchlaßbedingungen vorfinden. Da die Minoritätsträgerkonzentrationen bei entsprechender Dotierung sehr klein sind, kommt es durch Extraktion zu einem Absinken ihrer Dichten unter die Gleichgewichtswerte. Die Bezeichnungen Durchlaß- oder Sperrichtung beziehen sich demzufolge nur auf das Verhalten der Majoritätsträger. Im übrigen fallen bei gesperrten *pn*-Übergängen auch die Majoritätsträgerdichten innerhalb der eigentlichen Übergangszonen steil ab, denn für die hier vorhandenen Elektronen und Löcher ist der Durchtritt nach einer der beiden Seiten energetisch begünstigt. Insgesamt bildet sich an einem in Sperrichtung vorgespannten *pn*-Übergang demnach eine hochohmige Zone geringer Trägerdichten, die sogenannte Verarmungsschicht, aus.

Abschließend sei noch darauf hingewiesen, daß *pn*-Übergänge häufig bewußt unsymmetrisch ausgeführt werden, indem die eine Seite des Halbleiters um Größenordnungen stärker dotiert wird als die andere. Der Stromfluß durch den Übergang wird dann fast ausschließlich von den injizierten Majoritätsträgern aus der stärker dotierten Seite getragen, und auf der schwach dotierten Seite überwiegt der Stromanteil eingeschwemmter Minoritätsträger, solange der Verlust durch Rekombination gering bleibt. Dies ist vor allem für die Basiszonen von Bipolartransistoren wichtig (s. a. 8.3.6.3.).

8.3.6.2. Dioden und Leistungsgleichrichter. Wird an einen Halbleiter mit *pn*-Übergang eine Wechselspannung angelegt, so tritt ein Gleichrichtereffekt auf, da bei einer Polarität ein großer Durchlaßstrom, in der umgekehrten Richtung dagegen nur ein sehr kleiner Sperrstrom

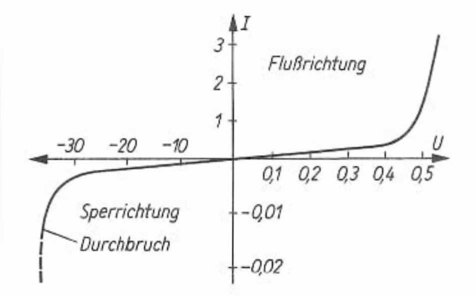


Abb. 8.3.6-2 Strom-Spannungs-Charakteristik einer Halbleiterdiode; die Abszisse ist unterschiedlich geteilt; Strom *I* in mA, Spannung *U* in V

fließt (Abb. 8.3.6-2). Bauelemente mit einem *pn*-Übergang und zwei Anschlüssen werden daher als *Dioden* in riesigen Stückzahlen produziert und eingesetzt; sie bilden darüber hinaus neben den Transistoren die wichtigsten Funktionseinheiten in den integrierten Schaltkreisen der Mikroelektronik. Die vielseitigen Anwendungen umfassen den Einsatz als Richtleiter in HF-Schaltungen, als Schalter in digitalen Logikschaltungen, als variable Kapazitäten, als Spannungsstabilisatoren und als Leistungsgleichrichter für hohe Ströme und Spannungen. Um die sehr unterschiedlichen Anforderungen zu erfüllen, werden die Fläche des *pn*-Übergangs, das Dotierungsprofil und andere den Kennlinienverlauf bestimmende Parameter dem jeweiligen Einsatz angepaßt. Nachfolgend werden einige wichtige Ausführungsformen charakterisiert (Tafel 45).

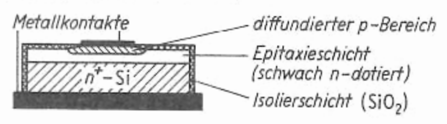


Abb. 8.3.6-3 Schematischer Aufbau einer diskreten Epitaxiediode mit $n+np$- bzw. *nip*-Übergang

8.3.6.2.1. *Spitzendioden.* Diese Bauelemente stellen Weiterentwicklungen des Bleiglanzdetektors aus der Anfangszeit der drahtlosen Nachrichtentechnik dar. Anstelle von Bleiglanz- (PbS-) Kristallen mit ihren wenig reproduzierbaren Eigenschaften benutzt man *n*-Ge-Kristalle, auf die beispielsweise eine Spitze von Golddraht mit etwas zulegiertem Gallium aufgedrückt und durch einen Stromimpuls angeschweißt wird. Dabei bildet sich örtlich eine *p*-dotierte Rekristallisationszone, also ein kleinflächiger *pn*-Übergang aus. Solche Dioden werden wegen ihrer kleinen Kapazität in begrenztem Umfang als Richtdioden im Mikrowellenbereich und als schnelle Schalter eingesetzt.

8.3.6.2.2. *Flächendioden für HF- und Impulsbetrieb.* Hinsichtlich der produzierten Stückzahl überwiegen bei weitem die Richt- und Schaltdioden mit flächenhaften, jedoch sehr kleinen *pn*-Übergängen. Man verwendet heute fast ausschließlich Silicium als Halbleiterwerkstoff, und die *pn*-Übergänge werden durch Eindiffundieren von Dotierungsmaterial, in geringerem Maße auch durch Ionenimplantation (s. a. 8.3.7.), erzeugt. Im Durchlaßbereich fließt Strom, wenn die angelegte Spannung einen Schwellwert von etwa 0,3 bis 0,7 V übersteigt. Der um Größenordnungen kleinere Sperrstrom ist stark temperaturabhängig; er steigt steil an, wenn die maximale Sperrspannung von einigen 10^1 bis 10^2 V überschritten wird, wobei infolge Durchbruchs eine Zerstörung der Diode ein-

treten kann. Die für elektronische Logikschaltungen, besonders in Mikroprozessoren und Computern, in sehr großer Stückzahl benötigten Schaltdioden sollen hohe Schaltgeschwindigkeit, große Durchlaßstromdichte bei minimaler Durchlaßspannung und einen möglichst kleinen Sperrstrom bei hoher Sperrspannung aufweisen. Diese Forderungen lassen sich am besten mit der *Epitaxietechnik* erfüllen. Abb. 8.3.6-3 zeigt schematisch eine diskrete Ausführungsform mit einem n^+-dotierten Si-Scheibchen als Grundmaterial. Hier soll das Pluszeichen auf eine hohe Dotantenkonzentration und somit einen geringen spezifischen Widerstand des Kristallplättchens hinweisen. Auf dieser niederohmigen Scheibe, die im wesentlichen nur als Träger dient, befindet sich eine hochohmige Epitaxieschicht, die durch Abscheiden von Si aus der Gasphase abgeschieden wird (s. a. 8.3.7.). In diese schwach n-dotierte Schicht läßt man aus der Gasphase von der Oberfläche her Akzeptoren eindiffundieren, wodurch ein pn-Übergang erzeugt wird. Wegen der nahezu eigenleitenden Zwischenschicht werden derartige Strukturen auch als *pin-Übergänge* bezeichnet. Um die Schaltgeschwindigkeit zu steigern, werden zusätzlich Rekombinationszentren in Form von Schwermetallatomen in die Übergangszone eingebaut. Natürlich läßt sich auch die umgekehrte Schichtenfolge mit p^+-leitendem Trägerkristall realisieren, und in vielen Halbleiterschaltungen wird die Möglichkeit zum Einsatz solcher komplementärer Bauelemente genutzt. Neben den diskreten Halbleiterdioden, die als Einzelbauelemente weiterhin benötigt werden, haben prinzipiell gleichartige Funktionseinheiten in stark miniaturisierter Ausführung von wenigen Mikrometern Durchmesser als Bestandteile integrierter Schaltkreise ständig an Bedeutung gewonnen.

8.3.6.2.3. *Kapazitätsdioden.* Gesperrte pn-Übergänge bestehen aus einer Verarmungszone mit isolierenden Eigenschaften, an die beidseitig die Raumladungszonen der p- und n-Bereiche grenzen (↑ 8.3.6.1.2.). Diese Struktur entspricht einem variablen Plattenkondensator, und die unvermeidliche Kapazität, die mit jedem pn-Übergang verbunden ist, begrenzt den Einsatz der Bauelemente bei sehr hohen Frequenzen im GHz-Bereich durch die klein werdende Impedanz, die zum HF-Kurzschluß führt. Auf der anderen Seite lassen sich pn-Übergänge in Form diskreter Kapazitätsdioden oder als Funktionseinheiten integrierter Schaltkreise für Filter, Modulationsschaltungen oder parametrische Verstärker einsetzen. Mit zunehmender Sperrspannung, aber unterhalb des Durchbruchs, nimmt die Kapazität ab, da sich die Raumladungs- und Verarmungszonen verbreitern.

8.3.6.2.4. *Tunneldioden.* Diese von ESAKI entwickelten Halbleiterbauelemente sind Dioden mit einem abrupten pn-Übergang von weniger als 10 nm Breite zwischen extrem hoch dotierten p- bzw. n-leitenden Bereichen. Die Majoritätsträgerkonzentrationen sind dann so hoch, daß das Ferminiveau einige k_BT von den Kanten entfernt innerhalb des Valenz- bzw. des Leitungsbands verläuft. Dies entspricht den Bedingungen entarteter Halbleiter (↑ 8.3.3.4.). Wegen des großen Konzentrationsgradienten wird am Übergang schon ohne angelegte Spannung eine relativ hohe Potentialdifferenz aufgebaut. Die Energiebänder werden dadurch so stark verschoben, daß sich das Valenzband auf der p-Seite mit dem Leitungsband auf der n-Seite teilweise überlappt (Abb. 8.3.6-4). Wird nun in Durchlaßrichtung eine kleine Vorspannung angelegt, so treten Elektronen aufgrund des quantenmechanischen Tunneleffekts von besetzten Zuständen des Leitungsbands auf der p-Seite in unbesetzte Valenzbandzustände auf der p-Seite über. Bei größerer Vorspannung wird die Bandüberlappung allerdings verschlechtert und verschwindet schließlich ganz.

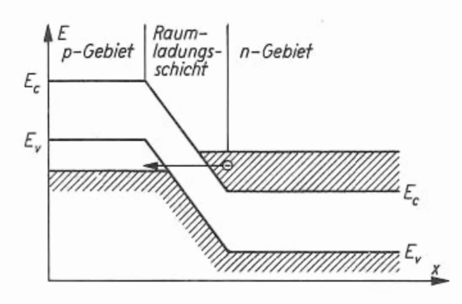


Abb. 8.3.6-4 Bandverlauf in der Nähe des pn-Übergangs einer Tunneldiode; die Raumladungsschicht darf nur wenige Nanometer dick sein

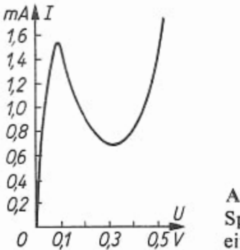


Abb. 8.3.6-5 Strom-Spannungs-Kennlinie einer Tunneldiode

Der Tunnelstrom nimmt daher nach Durchlaufen eines Maximums mit weiter wachsender Spannung ab, d. h. die Kennlinie weist einen Bereich mit fallender Charakteristik auf

(Abb. 8.3.6-5). Nach einem Minimum steigt der Diodenstrom wieder an, doch ist nun das Durchlaßverhalten normaler *pn*-Übergänge bestimmend. Aber auch bei umgekehrter Polung zeigen Tunneldioden kein Sperrverhalten, sondern es fließt ein mit zunehmender Spannung monoton und steil ansteigender Rückstrom. Verantwortlich hierfür ist der Schottkyeffekt, eine Art innerer Feldemission, die in der Erzeugung von Ladungsträgern durch feldinduzierte Band-Band-Übergänge unter dem Einfluß der sehr hohen elektrischen Feldstärke an steilen *pn*-Übergängen besteht.

Tunneldioden werden vorwiegend als Oszillatoren im Mikrowellenbereich bis zu mehreren 100 GHz, als parametrische Verstärker und als sehr schnelle Schalter eingesetzt.

8.3.6.2.5. *Zenerdioden.* Bei diesen auch als Z-Dioden bezeichneten Bauelementen wird der steile Stromanstieg im Durchbruchsbereich nach Überschreiten der maximalen Sperrspannung ausgenutzt. Mit Si als Halbleitermaterial lassen sich Z-Dioden mit Durchbruchsspannungen von etwa 5 bis 200 V bei maximaler Leistung bis zu 100 W herstellen. Neben dem eigentlichen Zenereffekt (↑ 8.3.5.5.) trägt auch die Ladungsträgervervielfachung bei einsetzendem Lawinendurchbruch zum Stromfluß in Z-Dioden bei. Falls der Zenereffekt überwiegt, ist der Temperaturkoeffizient der Durchbruchsspannung negativ, bei dominierendem Einfluß des Lawinendurchbruchs dagegen positiv. Durch Kombination von Z-Dioden mit positivem Temperaturgang und normalen Si-Dioden in Durchlaßrichtung mit negativem Temperaturgang lassen sich Referenzelemente für die Meßtechnik herstellen, die in einem bestimmten Temperaturgebiet eine nahezu konstante Charakteristik aufweisen. Im übrigen werden Z-Dioden hauptsächlich für die Spannungsbegrenzung, beispielsweise in Sicherheitsschaltern, eingesetzt.

8.3.6.2.6. *Leistungsgleichrichter.* Zur Gleichrichtung von Wechselstrom größerer Stromstärke und Spannung, beispielsweise für die Stromversorgung elektronischer Geräte, für Akkumulatorladegeräte, für Bahnanlagen sowie für elektrochemische Industrieanlagen, werden heute vorwiegend großflächige Si-Dioden eingesetzt, die dann als Halbleitergleichrichter bezeichnet werden. Das Verhalten wird wesentlich von der Perfektion großflächiger Einkristallscheiben (bis 150 mm Durchmesser), der gleichmäßigen Ausführung der *pn*-Übergänge und von einer optimalen Wärmeabfuhr über Metallunterlagen und Kühlrippen bestimmt. Si-Gleichrichter lassen sich für maximale Stromdichten von einigen 10^2 Acm^{-2} am Übergang und für Durchbruchsspannungen bis zu mehreren 10^3 V auslegen. Durch Parallel- bzw. Reihenschaltung solcher Einheiten können kompakte Gleichrichter für sehr hohe Ströme bzw. Spannungen hergestellt werden.

8.3.6.2.7. *Gesteuerte Gleichrichter – Thyristoren.* Darunter versteht man Leistungsbauelemente mit drei hintereinandergeschalteten *pn*-Übergängen in einer einkristallinen Siliciumstruktur,

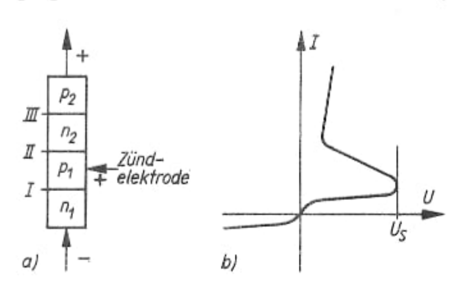


Abb. 8.3.6-6 a) Prinzipieller Aufbau und b) Schaltkennlinie eines Thyristors; U_s Schaltspannung

die häufig aus einer Kristallscheibe mit aufgebrachter Epitaxieschicht hergestellt wird. Wird an das System (Abb. 8.3.6-6) eine Spannung in der Weise angelegt, daß die *pn*-Übergänge I und III in Durchlaßrichtung gepolt sind, so kann zunächst kein nennenswerter Strom fließen, weil der Übergang II gesperrt ist. Da fast die Gesamtspannung an diesem Übergang abfällt, kommt es dort zum Lawinendurchbruch, sobald eine bestimmte Schaltspannung überschritten wird. Der Übergang wird dann mit Ladungsträgern überschwemmt, und das System wird leitend, wobei der Spannungsabfall im leitenden Zustand viel kleiner ist als die zur Einleitung des Durchbruchs erforderliche Schaltspannung. Derartige Gleichrichter müssen daher mit einem strombegrenzenden Vorwiderstand betrieben werden. Unterschreitet der Strom eine als Haltestrom bezeichnete Mindeststärke, so kippt das System wieder in den gesperrten Ausgangszustand zurück. Bei umgekehrter Polung erfolgt kein Schalten, sondern die maximale Sperrspannung ist hoch, weil die Übergänge I und III gesperrt sind.

Systeme dieser Art ohne zusätzliche Steuerung werden als Vierschichtdioden in Impulsschaltungen eingesetzt. Bei den steuerbaren Gleichrichtern oder Thyristoren ist am p_1-Gebiet ein metallischer Kontakt angebracht. Wird über diese Steuerelektrode ein Steuerstrom eingespeist, so verringert sich die Schaltspannung mit zunehmender Steuerstromstärke. Bei Steuerleistungen in der Größenordnung von 1 W und Schaltzeiten unter 1 µs lassen sich mit Thyristoren Ströme bis zu mehreren 10^2 A schalten. Die steuerbaren Gleichrichter finden in der Starkstromtechnik, insbesondere für Regelungsanlagen und elektrische Antriebe,

eine ausgedehnte Anwendung. Gegenüber herkömmlichen Schaltungen ermöglichen Thyristoren wesentliche Einsparungen an Elektroenergie, so daß ihr Einsatz zunehmend an ökonomischer Bedeutung gewinnt.

8.3.6.2.8. *Metall-Halbleiter-Sperrschichtdioden.* An Berührungsstellen zwischen Halbleitern und Metallen zeigen sich häufig Abweichungen

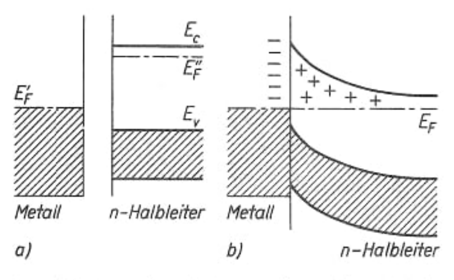


Abb. 8.3.6-7 Ausbildung einer Sperrschicht an einem Halbleiter-Metall-Kontakt; a) Bandverlauf vor bzw. b) nach der Berührung

gegenüber dem Leitungsverhalten entsprechend dem Ohmschen Gesetz, und es können sich Sperrschichten mit Gleichrichterverhalten ausbilden. Der Mechanismus dieser Erscheinungen, die wegen des Zusammenspiels verschiedener Volumen- und Grenzflächeneffekte ziemlich kompliziert sein können, wurde im wesentlichen Punkten von SCHOTTKY, MOTT und BARDEEN auf der Basis grundlegender Arbeiten von TAMM, JOFFÉ und FRENKEL aufgeklärt. Betrachtet sei der Fall eines n-Halbleiters in Kontakt mit einem Metall. Solange sich die Materialien nicht berühren, wird das Ferminiveau E_F' des Metalls von dem des Halbleiters E_F'' verschieden sein, und zwar fällt E_F für übliche Halbleiter-Metall-Kombinationen meist in die verbotene Zone zwischen Valenz- und Leitungsband des Halbleiters (Abb. 8.3.6-7). In diesem Fall stellt sich nach der Berührung das Gleichgewicht mit einheitlichem E_F ein, indem Elektronen an das Metall übertreten. Dadurch bildet sich eine elektrische Doppelschicht aus, die metallseitig von Elektronen an der Grenzfläche und auf der Halbleiterseite von der positiven Raumladung in einer an Ladungsträgern verarmten Grenzschicht bestimmt wird. Im Energiebandschema ist ein relatives Absinken der Bänder im Halbleitervolumen und eine positive Bandkrümmung im Bereich der Raumladungszone zu verzeichnen. Innerhalb der Raumladungszone liegt ein Mangel an Majoritätsträgern vor, der einen erhöhten Widerstand dieser Sperr- oder Verarmungsschicht bedingt. Wird nun eine äußere Spannung angelegt, so tritt Stromfluß ein, wenn die Halbleiterseite negativ, das Metall positiv gepolt ist, während bei umgekehrter Polarität nur ein geringer Reststrom von Minoritätsträgern beobachtet wird. Dieses Verhalten ist verständlich, weil die Bänder im Halbleitervolumen bei negativer Vorspannung angehoben werden und dadurch die Potentialschwelle zwischen Metall und Halbleiter verringert wird. Bei umgekehrter Polarität wird der sperrende Potentialwall dagegen vergrößert, und es kann erst bei hohen Spannungen infolge von Durchbruchserscheinungen oder bei sehr dünnen Randschichten auch durch Tunneleffekt ein wesentlicher Strom fließen. Der Fall eines p-Halbleiters in Kontakt mit einem Metall kann analog durch umgekehrten Elektronenübertritt vom Metall zum Halbleiter und eine negative Bandkrümmung in der Sperrschicht gedeutet werden. Flußrichtung liegt dann vor, wenn am Metall der negative und am Halbleiter der positive Pol angelegt wird. Allgemein gilt die Regel, daß bei Sperrschichtgleichrichtern Durchlaßverhalten auftritt, wenn die Majoritätsträger des Halbleiters zum Metall hin wandern.

Auf solchen Sperrschichten beruhen viele der früher benutzten Metallspitzendioden oder Detektoren (s. a. 8.3.6.2.1.) sowie die Selen- und Kupfer(I)-Oxidgleichrichter, die teilweise noch anzutreffen sind. Man stellt sie durch Aufdampfen von Selen auf Metalle bzw. durch oberflächliche Oxydation von Kupfer zu Cu_2O her. Als Schottkydioden werden neuere Ausführungen auf der Basis von Si, Ge, oder Verbindungshalbleitern mit punkt- oder flächenförmigen Metallkontakten bezeichnet. Trotz Schwierigkeiten hinsichtlich des reproduzierbaren Verhaltens haben sie neuerdings wegen der sich abzeichnenden Möglichkeiten einer extremen Miniaturisierung im Submikrometerbereich wieder ein gewisses Interesse erlangt. Unfreiwillige Dioden können sich an Kontakten von Halbleiterbauelementen ausbilden und bilden eine wichtige Ursache von Störungen. Um solche Effekte zu vermeiden, also Ohmsche Kontakte zu erhalten, muß man geeignete Materialpaarungen wählen und an der Grenzschicht des Halbleiters durch Aufrauhen oder Einlagerung von Störstellen für eine hohe Störstellendichte sorgen. Günstige Ergebnisse werden in der Mikroelektronik mit aufgedampften oder aufgestäubten Aluminiumkontakten erzielt. Das vorstehend entwickelte Modell der Halbleiter-Metall-Sperrschichten ist noch stark vereinfacht. Meist spielen zusätzlich spezielle Oberflächenzustände der Elektronen an der Halbleiteroberfläche oder innerhalb sehr dünner Oxidzwischenschichten eine wichtige Rolle. Ein weiteres interessantes Phänomen bilden die Inversionsschichten, die sich ausbilden, wenn die Bandverbiegung in der Raumladungszone so stark ist, daß der Leitungstyp vom n- zum p-Typ oder umgekehrt umschlägt.

Dadurch entsteht ein sehr schmaler *pn*-Übergang, eine Erscheinung, die bei den älteren Metallspitzendioden und -transistoren unbewußt oder gezielt genutzt wurde.

8.3.6.2.9. *Foto- und Lumineszenzdioden.* Der Vollständigkeit halber sind hier diese wichtigen Bauelemente der *Optoelektronik* zu erwähnen, die in einem späteren Abschnitt ausführlicher behandelt werden (↑ 8.8.3.). Fotodioden liefern bei Lichteinfall eine elektrische Spannung, die Foto-EMK, die für meßtechnische Zwecke oder bei größeren Einheiten zur Energiegewinnung mittels Solarzellen genutzt wird. Lumineszenzdioden senden bei Stromfluß Licht aus; sie werden für die Signalanzeige bei Rechnern, Uhren und den verschiedensten Meß- oder Kontrollgeräten in immer größer werdendem Ausmaß eingesetzt.

8.3.6.3. Bipolare Transistoren. Die Entwicklung des Spitzentransistors durch Bardeen und Brattain im Jahre 1947, die wenig später folgende Konzipierung des Flächentransistors durch Shockley und die Anfang der 50er Jahre einsetzende Serienproduktion dieser Bauelemente sind als Marksteine in der Geschichte der Halbleitertechnik anzusehen. Als Transistor im engeren Sinne wird heute der bipolare Flächentransistor verstanden, und die Kennzeichnung „bipolar" kann weggelassen werden, solange nicht die Möglichkeit einer Verwechslung mit den unipolaren *Feldeffekt-* oder *MOS-Transistoren* (↑ 8.3.6.4.) gegeben ist. Ein Flächentransistor besteht aus einer Halbleiterstruktur mit drei verschieden dotierten Kristallbereichen, die durch zwei *pn*-Übergänge voneinander getrennt sind. Die drei Bereiche, die mit getrennten metallischen Kontakten versehen sind, werden Emitter (E), Basis (B) und Kollektor (C) genannt (Abb. 8.3.6-8). Sind Emitter- und Kollektorbereich *n*-dotiert, die dazwischenliegende Basis *p*-dotiert, so spricht man von einem *npn*-Transistor. Das komplementäre Bauelement mit umgekehrt dotierten Bereichen heißt *pnp*-Transistor. Beide Typen verhalten sich qualitativ gleichartig, wenn auch die anliegenden Spannungen entgegengesetztes Vorzeichen tragen.

8.3.6.3.1. *Transistor in Emitterschaltung.* Die Funktionsweise soll am Beispiel eines *npn*-Transistors in Emitterschaltung diskutiert werden (Abb. 8.3.6-9). Der Emitter, der in diesem Fall meist geerdet wird, gehört bei dieser Schaltvariante gemeinsam dem Eingangs- und dem Ausgangsstromkreis an. Angenommen sei nun, daß der Emitter auf dem Massepotential Null gehalten wird, während an der Basis eine kleine und am Kollektor eine größere positive Spannung anliegt. In der Praxis beträgt die Betriebsspannung von Transistorgeräten meist 6 oder 9 V. Zerlegt man den Transistor gedanklich in eine Emitter-Basis-(EB-) und eine Basis-Kollektor-(BC)-Diode, so ist die erste bei den angenommenen Spannungsverhältnissen in Durchlaßrichtung gepolt, die zweite dagegen gesperrt. Das bezieht sich freilich nur auf die jeweiligen Majoritätsträger. Als Folge des Stromflusses in der EB-Diode werden Elektronen in die Basiszone eingeschwemmt, und durch diese Injektion wird die Konzentration der Minoritätsträger im Basisbereich wesentlich über die Gleichgewichtsdichte angehoben. Natürlich findet umgekehrt auch eine Injektion von Löchern aus der Basis in den Emitterbereich statt. Durch stärkere Dotierung der Emitterzone im Vergleich zum Basisgebiet läßt sich jedoch erreichen, daß der Strom durch den EB-Übergang überwiegend durch die vom Emitter in die Basis injizierten Ladungsträger bestimmt wird. Als *Emitterwirkungsgrad* γ definiert man den Anteil des in die Basis injizierten Stroms am Gesamtstrom I_{EB} durch den EB-Übergang; er kann mehr als 0,99 erreichen.

Für die Funktion des Transistors als aktives elektronisches Bauelement mit Verstärkereigenschaften ist es weiterhin entscheidend, daß ein möglichst hoher Anteil der in die Basis eingeschwemmten Überschußladungsträger (im gewählten Beispiel Elektronen) durch Diffusion bis zum BC-Übergang gelangt. Dort finden sie als Minoritätsträger an dem für Majoritätsträger gesperrten Übergang gerade Durchlaßbedingungen vor (s. a. 8.3.6.1.2.). Sie treten daher mit hoher Wahrscheinlichkeit in den



Abb. 8.3.6-8 a) Prinzipieller Aufbau, b) Schaltsymbol eines (bipolaren) Transistors

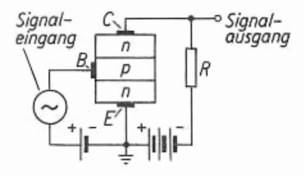


Abb. 8.3.6-9 *npn*-Transistor in Ermittlungsschaltung

Kollektorbereich über, wo sie wieder zu Majoritätsträgern werden und unter dem Einfluß des elektrischen Felds zur Kollektorelektrode wandern. Der *Transport-* oder *Einfangfaktor* β, welcher den Anteil der in die Basis injizierten Träger angibt, der tatsächlich die Kollektorzone erreicht, liegt gewöhnlich zwischen 0,95

und 0,999. Der Kollektorbereich ist meist stärker als die Basiszone, jedoch schwächer als die Emitterzone dotiert. Um die Verluste in der Basiszone durch Volumen- bzw. Oberflächenrekombination (↑ 8.3.4.2.) möglichst klein zu halten, soll die Basis möglichst dünn sein, bei hoher kristalliner Perfektion eine geringe Dichte von Rekombinationszentren enthalten und eine minimale Oberfläche besitzen.

Das Produkt $\alpha = \beta\gamma$ aus Transportfaktor und Emitterwirkungsgrad wird *Stromverstärkungsfaktor* genannt und ist stets kleiner als 1. Trotzdem ist mit Transistoren natürlich eine Stromverstärkung möglich, wie sich aus der folgenden Überlegung für die Emitterschaltung ergibt. Ist I_{EB} die Stromstärke am EB-Übergang, so beträgt die Stromstärke am BC-Übergang $I_{BC} = I_0 + \alpha I_{EB}$, wobei mit I_0 der kleine Reststrom berücksichtigt wurde, der nicht von den aus dem Emitter injizierten Ladungsträgern gestellt wird. In der Zuleitung zur Basis fließt dann der Strom $I_B = I_{EB} - I_{BC} = (1 - \alpha)\, I_{EB} - I_0$, der als Differenz sehr viel kleiner ist als der Stromfluß am Emitter oder Kollektor.

Dieser Basisstrom wird nun verändert, sobald im Eingangsstromkreis elektrische Signale ankommen. Dies wiederum hat steile Änderungen der Potentialdifferenz am EB-Übergang zur Folge, die sich exponentiell auf den Durchlaßstrom I_{EB} und damit auf den Stromfluß im Ausgangskreis auswirken. Durch kleine Stromänderungen an der Basis wird der viel größere Ausgangsstrom im Emitter-Kollektor-Stromkreis daher in empfindlicher Weise gesteuert, wobei die Stromverstärkung das 10- bis 100-fache erreichen kann. Daneben läßt sich eine Spannungsverstärkung erzielen, deren Größe durch das Verhältnis der Wechselstromwiderstände oder Impedanzen in den Teilstromkreisen bestimmt wird. Es sei aber ausdrücklich betont, daß die Steuerung beim Transistor wegen der relativ niedrigen Eingangsimpedanz nicht in erster Linie elektrostatisch, sondern über einen Stromfluß vor sich geht. Bipolare Transistoren sind im Unterschied zu Elektronenröhren oder MOS-Transistoren stromgesteuerte Bauelemente.

Das genaue Verstärkungsverhalten läßt sich unter Benutzung der Kennliniendaten ermitteln, die man aus Messungen gewinnt. Zumindest qualitativ kann der Kennlinienverlauf jedoch auch theoretisch mit Hilfe der Trägertransportgleichungen 8.3.(6) berechnet werden. Im Fall der Emitterschaltung folgt, daß Eingangs- und Ausgangssignal gegenphasig sind. Eine Strom- bzw. Spannungszunahme am Eingang bewirkt eine Abnahme des Stroms bzw. der Spannung am Ausgang; die Signale werden also umgekehrt.

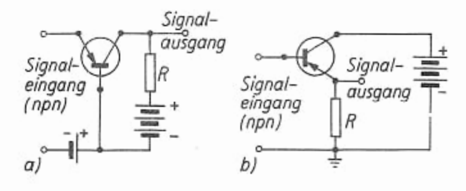


Abb. 8.3.6-10 *npn*-Transistor a) in Basis- und b) in Kollektorschaltung

8.3.6.3.2. *Basis- und Kollektorschaltung.* Eine weitere wichtige Schaltvariante bildet die Basisschaltung (Abb. 8.3.6-10a), bei der die Basis dem Eingangs- und dem Ausgangsstromkreis gemeinsam angehört. Für die in den Emitter eintretenden Signale ist die Stromverstärkung entsprechend dem oben eingeführten Faktor α stets kleiner als 1, aber dafür läßt sich eine hohe Spannungsverstärkung erzielen. Bei der seltener angewendeten Schaltung mit gemeinsamem Kollektor in Eingangs- und Ausgangskreis (Abb. 8.3.6-10b) ist schließlich die Spannungsverstärkung kleiner als 1, während eine hohe Stromverstärkung möglich ist, die beispielsweise für die Verstärkung leistungsschwacher Signale in Vorverstärkern von Antennen benutzt wird.

8.3.6.3.3. *Diskrete Transistorbauelemente.* Diskrete *npn*- und *pnp*-Transistoren werden in einem breiten Sortiment mit unterschiedlichen Parametern bezüglich des Kennlinienverlaufs, des Hochfrequenzverhaltens oder der maximalen Leistung hergestellt. Als Material findet fast ausschließlich Silicium Einsatz; nur für sehr hohe Frequenzen wird noch Germanium verwendet. Der Einsatz von $A^{III}B^{V}$-Verbindungshalbleitern wie GaAs, GaP oder $GaAs_xP_{1-x}$ beschränkt sich im wesentlichen auf Fototransistoren (s. a. 8.8.3.), bei denen der Bereich des EB-Übergangs gleichzeitig als Foto- oder als Lumineszenzdiode wirkt.

Die Einarbeitung der verschieden dotierten Bereiche geschieht überwiegend in Form der Planartechnologie von einer Oberfläche der Kristallscheibe her. Die aus einem homogen dotierten Einkristall herausgeschnittenen dünnen Si-Scheiben werden durch genau kontrollierte Behandlung in einer Sauerstoffatmosphäre bei Temperaturen um 1200 °C oberflächlich oxydiert, wobei sich eine SiO_2-Schicht bildet. In diese Schicht werden mittels einer Fotomaskierungstechnik (↑ 8.3.7.) und anschließender Ätzung Fenster eingearbeitet, durch die man in einem weiteren Hochtemperaturprozeß Atome des Dotierungsmaterials eindiffundieren läßt. Die Scheiben werden dazu in Quarzrohröfen geeigneten, beispielsweise bor- oder phosphorhaltigen, Dämpfen ausgesetzt. Durch sehr genaue Regelung der Verfahrensparameter, insbesondere der Temperatur, läßt sich erreichen, daß die Tiefe der umdotierten Be-

reiche mit einer Genauigkeit von größenordnungsmäßig 1 μm definiert werden kann und zugleich eine steile Diffusionsfront entsteht. Nach Herstellung der Basisgebiete im Grundmaterial, das gewöhnlich als Kollektorbereich verwendet wird, müssen diese Verfahrensschritte nochmals bei veränderter Geometrie wiederholt werden, um bei erneuter Umdotierung die Emitterzonen einzuarbeiten. Die Kontaktierung erfolgt durch selektives Aufdampfen oder Aufstäuben von Metallen (Al, Cr, Au) unter Benutzung geeigneter Metallmasken. Die Anschlüsse werden durch Bonden, d. h. Anschweißen von Drähten, hergestellt, die bei HF-Transistoren aus Gold bestehen. Die fertigen Elemente werden mit einer schützenden Schicht versehen und in Metall- oder Plastgehäuse verkapselt. Eine wichtige Variante ist die Epitaxietechnik (↑ 8.3.6.2.2.). Bei dieser Technologie läßt man auf den Kristallscheiben zunächst eine Si-Schicht aus der Gasphase epitaktisch, d. h. einkristallin bei gleicher Orientierung wie die Unterlage, aufwachsen. In dieser Schicht werden mit den vorstehend umrissenen Prozeßschritten die verschieden dotierten Bereiche erzeugt. Damit läßt sich beispielsweise erreichen, daß der *pn*-Übergang zwischen Kollektor und Basis eine dünne, nahezu eigenleitende Zwischenschicht enthält. Man spricht dann von einem *pin-Übergang* (↑ 8.3.6.2.2.). Als Vorteile ergeben sich kleinere Sperrschichtkapazitäten und somit ein verbessertes Hochfrequenzverhalten sowie höhere Durchbruchspannungen.

Eine hohe Produktivität bei der Herstellung von Transistoren und anderen Halbleiterbauelementen wird dadurch erzielt, daß bei den grundlegenden Verfahrensschritten Maskierung, Ätzung, Diffusion und Aufbringen der Kontakte gleichzeitig eine große Anzahl von Strukturen erzeugt wird. So lassen sich aus einer Si-Scheibe mit 100 mm Durchmesser und einigen Zehntel mm Dicke mehrere hundert bis tausend diskrete Halbleiterbauelemente herstellen. Die Bearbeitung einer Charge solcher Scheiben geschieht in der Weise, daß zunächst die diffundierten und kontaktierten Bereiche für viele Bauelemente erzeugt werden und erst danach eine Aufteilung der Scheiben in einzelne „Chips" vorgenommen wird. Für Hochfrequenztransistoren sind kleine Abmessungen nicht nur aus Gründen des sparsamen Materialeinsatzes erwünscht, sondern auch physikalisch notwendig, weil großflächige *pn*-Übergänge eine höhere Sperrschichtkapazität aufweisen, wodurch die Funktion bei hohen Frequenzen begrenzt wird. Das Haupteinsatzgebiet für diskrete HF- und Schalttransistoren bilden heute solche Anwendungsfälle, bei denen man mit einer geringen Anzahl von aktiven Bauelementen auskommt. Allerdings geht man auch bei diesen Anwendungen in zunehmendem Maße dazu über, aktive und passive Bauelemente zu standardisierten Funktionseinheiten zusammenzufassen, die in hohen Stückzahlen gefertigt werden. Bei dieser *Hybridtechnik* wird die passive Schaltung, die vorwiegend Widerstände und Leiterstränge umfaßt, in Form dünner Schichten auf Glas- oder Keramiksubstraten von einigen cm² Fläche abgeschieden. Die diskreten aktiven Bauelemente werden an diese Schaltungen angeschweißt oder -gelötet, und die so erhaltenen *Hybridschaltungen* verkapselt man in Plast- oder Keramikgehäusen. In Einsatzgebieten, die eine große Anzahl von aktiven Bauelementen erfordern, besonders bei den Logikschaltungen der elektronischen Rechen- und Steuerungstechnik, sind die diskreten Transistoren dagegen weitgehend durch die integrierten Schaltungen der Mikroelektronik verdrängt worden. Unentbehrlich sind diskrete Transistoren auch künftig bei den Leistungsstufen in den Ausgängen von elektronischen Geräten der Industrie- und Konsumelektronik. Für Leistungstransistoren benutzt man größere Si-Chips mit *pn*-Übergängen entsprechend größerer Fläche. Eine hohe Betriebsspannung läßt sich mit *pin*-Übergängen erreichen. Die maximale Betriebsleistung liegt üblicherweise zwischen 1 und 100 W, jedoch werden auch Einheiten mit einer Belastbarkeit von größenordnungsmäßig 1 kW gefertigt.

8.3.6.4. MOS-Transistoren. Eine wichtige Alternative zur bipolaren Halbleitertechnik bilden die *unipolaren Feldeffekttransistoren*, bei denen der Strom von Majoritätsträgern in einem schmalen Leitungskanal nahe der Halbleiteroberfläche vorwiegend elektrostatisch gesteuert wird. Nach der für diese Bauelemente typischen Schichtenfolge *M*etall/*O*xid/Halbleiter [*semiconductor* engl., Halbleiter] bezeichnet man sie als *MOS-Feldeffekttransistoren*, abgekürzt MOS-FET, oder einfach als MOS-Transistoren. Entsprechend wird die auf solchen aktiven Elementen basierende Variante der Mikroelektronik *MOS-Technik* genannt. Da die Oxidschicht auch durch andere Isolatorschichten, beispielsweise Nitrid, ersetzt werden kann, spricht man auch von *MIS-Transistoren* bzw. einer *MIS-Technik*.

8.3.6.4.1. *Feldeffekt.* Die Funktion von MOS-Transistoren beruht auf der als Feldeffekt seit langem bekannten Tatsache, wonach die Leitfähigkeit einer oberflächennahen Zone im Halbleiter durch äußere elektrische Felder in weiten Grenzen beeinflußt werden kann. Das Auftreten von Bandverbiegungen, die zur Bildung von *Inversionsschichten* führen können, wurde bereits im Zusammenhang mit Halbleiter/Metallkontakten besprochen (↑ 8.3.6.2.9.). In gezielter Weise läßt sich der Feldeffekt beobachten, wenn zwischen zwei dichtbenachbar-

ten Kontakten an einer Halbleiteroberfläche eine Elektrode angebracht wird, die vom Halbleiter durch eine dünne Isolatorschicht getrennt ist. Beim Anlegen einer entsprechend gepolten Spannung an diese Elektrode wird eine Inversionsschicht entgegengesetzten Leitungstyps erzeugt, deren Breite und Leitfähigkeit mit wachsender Steuerspannung zunehmen. Dadurch bildet sich eine *Anreicherungsschicht* aus, und der Stromfluß zwischen den beiden Kontakten läßt sich durch Variation der Steuerspannung empfindlich beeinflussen. Im Falle umgekehrter Polung kommt es in der Oberflächenzone zu einer *Verarmung* von Ladungsträgern. Das läßt sich gleichfalls zur Steuerung des Stroms ausnutzen, falls parallel zur Halbleiteroberfläche durch passende Dotierung ein *pn*-Übergang geschaffen wird, der den stromführenden Kanal vom Halbleitervolumen trennt. Seinem Wesen nach beruht der Feldeffekt auf elektrostatischer Wechselwirkung, und er kann als Auswirkung der Influenz aufgefaßt werden. Da die SiO_2-Schichten zwischen Halbleiter und Steuerelektrode einen sehr hohen Isolationswiderstand besitzen, werden zur Steuerung nur sehr kleine Ströme bzw. Leistungen benötigt. Im Gegensatz zur Stromsteuerung des bipolaren Transistors liegt also ausgesprochene Spannungssteuerung vor, und die MOS-Transistoren ähneln in dieser Hinsicht den Elektronenröhren. Allerdings wird bei MOS-Transistoren eine größere Steuerspannung von 4 bis 6 V benötigt, während Bipolartransistoren schon bei Eingangsspannungen von 0,5 bis 1 V kräftige Ausgangsströme liefern. Im übrigen hängen wichtige Parameter, wie die Steuerspannung oder die Schaltgeschwindigkeit, bei MOS-Transistoren stark von den *Oberflächenzuständen der Elektronen* ab, die sich als Folge von Störungen der Festkörperstruktur an der Grenzfläche Halbleiter/Isolator ausbilden. Diese Oberflächenzustände müssen jeweils erst gefüllt oder entleert werden, bevor das steuernde elektrische Feld voll auf den Halbleiter einwirken kann.

8.3.6.4.2. *Grundtypen von MOS-Transistoren.* Je nachdem, ob der Strom im Leitungskanal von Elektronen oder von Löchern als Majoritätsträger transportiert wird, spricht man von *n*- oder *p-Kanal-MOS-Transistoren.* Beide Varianten lassen sich als *Feldeffekttransistoren* im *Anreicherungs-* oder im *Verarmungsbetrieb* ausführen, so daß vier Grundtypen von MOS-Transistoren zu unterscheiden sind. Anhand von Abb. 8.3.6-11 betrachten wir den prinzipiellen Aufbau und die Funktion von *n*-Kanal-MOS-Transistoren der beiden Betriebsweisen. Bei der fast ausschließlich verwendeten *Silicium-Planar-Technologie* werden mit Hilfe der Diffusionstechnik im *p*-dotierten Si-Substrat zwei streifenförmige *n*-dotierte Oberflächenbereiche geschaffen. Im Falle des Anreicherungsbetriebs sind diese umdotierten Inseln voneinander getrennt, während bei der Verarmungsvariante eine *n*-leitende Brücke hergestellt wird. Die Breite des Spalts zwischen den beiden Inseln beträgt gewöhnlich nur einige µm; bei der hochminiaturisierten Kurzkanaltechnik ist sie sogar kleiner als 1 µm. Die Länge der Inseln bzw. des zwischen ihnen befindlichen Kanals richtet sich nach der gewünschten Leistung. Bei HF- und Schalttransistoren genügen Mikrometer während MOS-Leistungstransistoren Kanallängen von mehreren oder vielen Millimetern haben.

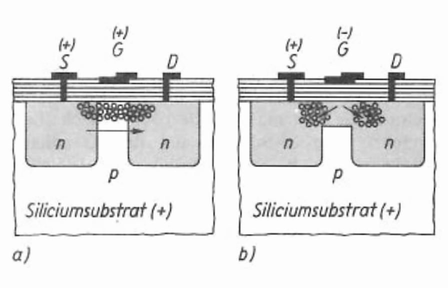


Abb. 8.3.6-11 *n*-Kanal-MOS-Transistoren; a) im Anreicherungs- b) im Verarmungsbetrieb

Ebenso wie bei der Bipolartechnik werden gleichzeitig viele Bauelemente in dünne Scheiben von Siliciumkristallen eingearbeitet, die später entweder in diskrete Bauelemente zerlegt oder als integrierte MOS-Schaltkreise gestaltet werden. Durch thermische Oxydation, Maskieren, Ätzen und selektives Metallaufdampfen stellt man die dargestellten Strukturen her. Wegen der kleinen Abmessungen ist dabei höchste Genauigkeit erforderlich. Insbesondere kommt es darauf an, daß die vom Halbleiter und den beiden anderen Kontakten isolierte Steuerelektrode genau über der Mitte des Kanalspalts positioniert ist. Die drei Elektroden des Bauelements werden üblicherweise mit S [*source* engl., Quelle], D [*drain* engl., Senke] und G [*gate* engl., Tor, Schranke] bezeichnet. Betrachtet sei nun zunächst der *Anreicherungsbetrieb* (↑ Abb. 8.3.6-11a). Wird zwischen S und D eine Betriebsspannung von einigen Volt angelegt, so kann zunächst kein nennenswerter Strom fließen, weil im p-dotierten Kanal nur wenige Elektronen vorhanden sind. Erst bei ausreichender positiver Vorspannung der Steuerelektrode G bildet sich eine *n*-leitende Inversionsschicht aus. Die Stärke des dann fließenden Stroms hängt empfindlich von der an G anliegenden Spannung ab. Auf diese Weise wird durch leistungsarme Steuerung des Stroms zwischen S und D die Verstärkung eines an G erscheinenden Spannungssignals

bewirkt. Im Falle des *Verarmungsbetriebs* (↑ Abb. 8.3.6-11 b) fließt bereits ohne Steuerspannung ein Strom über die *n*-dotierte Brücke. Wird nun eine negative Steuerspannung an G angelegt, so werden die Elektronen aus dem Kanal durch elektrostatische Abstoßung verdrängt. Der Stromfluß nimmt also mit zunehmendem Betrag der Steuerspannung ab, woraus gleichfalls ein Verstärkerverhalten resultiert. Im übrigen ist bei beiden Betriebsformen dafür zu sorgen, daß die *n*-dotierten Bereiche durch eine Vorspannung in Sperrichtung vom *p*-leitenden Grundgebiet des Kristalls isoliert werden.

8.3.6.4.3. *C-MOS-Technik.* MOS-Transistoren, die zu den beschriebenen Typen mit *n*-Kanal komplementär sind, lassen sich analog herstellen, indem *p*-dotierte Inseln in einer *n*-dotierten Kristallscheibe geschaffen werden. Beim Betrieb sind die Spannungen umzukehren. Auf der Kombination von *n*- und *p*-Kanal-MOS-Transistoren beruht die C-MOS-Technik. Vereinigt man nämlich komplementäre MOS-Transistoren zu logischen Schaltungen, so fließt im Mittel nur durch jedes zweite Funktionselement Strom, während die andere Hälfte der Bauelemente gesperrt ist. Dadurch verringert sich der Stromverbrauch, was für die batteriegespeisten Schaltungen von Taschenrechnern und Quarzuhren wichtig ist. Solche Schaltungen werden in integrierter Form in Halbleiterchips eingearbeitet, wobei gegenüber den diskreten Elementen ein weiterer Diffusionsschritt erforderlich ist, um komplementäre MOS-Funktionselemente auf dem gleichen Substratkristall schaffen zu können.

8.3.6.4.4. *Neuere Entwicklungen.* Infolge des einfachen Grundprinzips und der extrem hohen Miniaturisierungsmöglichkeit hat die MOS-Technik in Verbindung mit der Mikroelektronik eine ständig wachsende Bedeutung erlangt, ohne daß jedoch mit einer Verdrängung der Bipolartechnik zu rechnen wäre. Auf einige neuere Entwicklungen soll nachfolgend stichwortartig hingewiesen werden: Bei der *Si-Gate-Technik* werden die Steuerelektroden nicht durch Aufdampfen von Metallen wie Al, sondern durch lokale Abscheidung von dotiertem, polykristallinem Si hergestellt. Das bringt technologisch sowie hinsichtlich der Betriebsparameter Vorteile.

Durch Einsatz der *Ionenimplantation* (↑ 8.3.3.1.) in Verbindung mit dem Ausheilen von Gitterstörungen durch Wärmebehandlung gelingt es, die Genauigkeit bei der Schaffung der umdotierten Inseln bzw. Kanäle zu verbessern und die Konzentration an störenden Oberflächenzuständen der Elektronen herabzusetzen.

Bei den *Schottky-Feldeffekttransistoren*, deren Entwicklung im Fluß ist, wird zwischen Halbleiter und Steuerelektrode auf eine isolierende Zwischenschicht verzichtet und stattdessen die Bandverbiegung am Halbleiter/Metallkontakt ausgenutzt. Infolge höherer Perfektion der Grenzflächenstruktur wird die Ladungsträgerbeweglichkeit im Kanal größer, wodurch der Betrieb bei sehr hohen Frequenzen (bis 10^2 GHz) und ein entsprechend schnelles Schaltverhalten bei noch weitergehender Miniaturisierung realisierbar erscheinen.

Gleichfalls auf höchste Schaltgeschwindigkeiten bei extremer Miniaturisierung zielen Bemühungen ab, anstelle von Si Schichten aus GaAs oder anderen *$A^{III}B^{V}$-Halbleitern* zu verwenden. Um die ohnedies hohen Elektronenbeweglichkeiten dieser Materialien zu erhöhen, erwägt man den Betrieb bei tiefen Temperaturen. Die eigentlichen Kanalbereiche dürfen dann auch nicht dotiert sein, weil die ionisierten Störstellen die Beweglichkeit herabsetzen (s. a. 8.3.5.3.). Die für den Stromtransport erforderlichen Ladungsträger werden durch Eindiffusion oder Injektion aus benachbarten, dotierten Halbleiterschichten erhalten. Zur Herstellung der hierfür erforderlichen epitaktischen Schichtstrukturen verwendet man die *Molekularstrahltechnik*, bei der die Materialien unter Ultrahochvakuumbedingungen durch Kondensation definierter Dampfströme abgeschieden werden. Extrem hohe Schaltgeschwindigkeiten im Pikosekundenbereich, die mit solchen Anordnungen versuchsweise realisiert werden konnten, werden als *„ballistischer Elektronentransport"* von „heißen" Ladungsträgern gedeutet.

Technisch bereits praktiziert wird die *SOS-Technik*, bei der die Halbleiterschicht, meist Si, auf isolierenden Substraten abgeschieden wird. Zum Unterschied von der Homoepitaxie von gleichartigen Materialien, beispielsweise Si auf Si, spricht man von *Heteroepitaxie*, wenn eine Schicht einkristallin auf einem andersartigen Material aufwächst. Als isolierende, einkristalline Substrate für Si sind vor allem künstliche Halbedelsteine wie Saphir oder Spinell geeignet. Leider sind weitgehend perfekte Substrate dieser Art vorläufig zu teuer, doch zeichnet sich mit der Entwicklung der kontinuierlichen Bändchenkristallziehtechnik für Saphir ein Wandel ab. Besondere Vorteile der SOS-Technik leiten sich aus der Tatsache ab, daß die störenden Sperrschichtkapazitäten zwischen Funktionselementen und Substraten drastisch verringert werden können.

Eine neuartige Bauelementekonzeption, die aus der MOS-Technik hervorgegangen ist, bildet die *CCD-Technik* [*charge coupled devices* engl., ladungsgekoppelte Anordnungen]. Eine integrierte Schaltung dieser Art (Abb. 8.3.6-12) kann als ein extrem gestreckter MOS-Transistor aufgefaßt werden, der über dem Leitungskanal eine große Anzahl (bis zu 10^3) von Steuerelektroden besitzt. Bei entsprechender

Ansteuerung mit elektrischen Impulsen bilden sich Pakete von jeweils einigen hundert Ladungsträgern, die von Elektrode zu Elektrode wandern. Im Betrieb zirkuliert die Kette dieser Ladungspakete ständig durch die Anordnung.

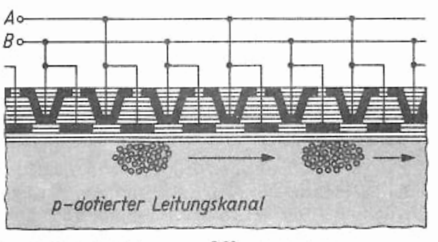


Abb. 8.3.6-12 Ausschnitt aus einer CCD-Anordnung; zeitlich versetzte Impulsfolgen an den Anschlüssen *A* und *B* bewegen die „Ladungspakete" von einer Elektrode zur benachbarten

Derartige Bauelemente lassen sich als schnelle und hochintegrierte Informationsspeicher einsetzen, indem man die An- oder Abwesenheit eines Ladungspakets an einer bestimmten Stelle der Kette den Ziffern 0 und L eines binären Zahlensystems zuordnet.

MOS-Transistoren mit freiliegender Oberfläche der Isolatorschicht zwischen Halbleiter und Steuerelektrode werden in ihrem elektrischen Verhalten durch die Zusammensetzung der umgebenden Atmosphäre beeinflußt. Dies ist auf das Eindiffundieren von Gasen und die dadurch bedingte Veränderung der lokalisierten Elektronenzustände in der Schicht und an der Halbleiter/Metallgrenzfläche zurückzuführen. Durch Einsatz spezieller Schichtstrukturen ist eine selektive Reaktion gegenüber bestimmten Gasen und Dämpfen möglich. Darauf beruhen Ansätze zur Entwicklung von *Gasdetektoren*, die als „elektronische Nasen" für Sensoren in der Automatisierungs- und Industrierobotertechnik wichtige Anwendungen finden könnten.

8.3.7. Mikroelektronik

8.3.7.1. Schaltungsintegration auf der Basis von Bipolar- und MOS-Technik. Die Schaltungen der klassischen Röhrenelektronik waren aus diskreten Bauelementen zusammengesetzt. Als *passive Bauelemente* verwendete man Widerstände, Kondensatoren, Induktivitäten oder Spulen sowie Kontaktbauelemente wie Schalter oder Steckverbinder. Zu den *aktiven Bauelementen* rechnete man Elektronenröhren mit Verstärker- oder zumindest Gleichrichtereigenschaften. Mit dem Einzug der Halbleitertechnik in die Elektronik am Anfang der 60er Jahre wurde die Konzeption diskreter Schaltungen zunächst übernommen, indem Transistoren bzw. Halbleiterdioden in die Schaltungen eingesetzt wurden. Eine Miniaturisierung war dabei vor allem wegen der kleineren Betriebsspannungen und des verringerten Leistungsbedarfs möglich. Auch heute spielen Schaltungen, die aus diskreten Bauelementen zusammengesetzt sind, für kleinere Geräte, für Netzversorgungseinheiten sowie Leistungsstufen und allgemein für die Peripherie elektronischer Anlagen eine wichtige Rolle, und daran wird sich in absehbarer Zukunft nichts ändern. Trotzdem wird der extrem steile Zuwachs im Produktionsvolumen und dem Einsatz elektronischer Schaltungen, der sich gegenwärtig in den Industriestaaten vollzieht, hauptsächlich durch die Schaltungsintegration entsprechend dem Konzept der Mikroelektronik bestimmt und überhaupt erst ermöglicht. Der Leitgedanke bei der Entwicklung der halbleiterphysikalischen Basis der Mikroelektronik besteht in der Schaffung kompletter elektronischer Schaltungen auf einem einzigen Siliciumplättchen von einigen mm Kantenlänge. Richtungweisend hierfür war die Erkenntnis, daß sich an der Oberfläche des Si-Chips viele nebeneinanderliegende Bereiche mit Transistor- oder Diodenfunktion erzeugen lassen, die man nur noch mit Hilfe eines Netzwerks aufgedampfter metallischer Leiterbahnen zu verbinden braucht, um eine fertige Schaltung zu erhalten. Noch vor 1970 wurden die ersten *integrierten Schaltkreise*, auch *Festkörperschaltkreise* genannt, in Verwirklichung dieses Konzepts serienmäßig gefertigt. Damit wurde eine qualitativ neue Entwicklung der Elektronik und Informationstechnik eingeleitet, deren Auswirkungen in immer stärkerem Maße das Niveau der gesamten Volkswirtschaft mitbestimmen.

Physikalisch beruht die Funktion der Festkörperschaltkreise auf den gleichen Prinzipien wie die der diskreten Halbleiterbauelemente, besonders der Bipolar- oder MOS-Transistoren. Zu beachten ist dabei, daß eine Vermischung der Bipolar- bzw. MOS-Technik wegen der unterschiedlichen Betriebsparameter normalerweise nicht in Betracht kommt. Ein integrierter Schaltkreis wird demgemäß entweder nach dem Bipolar- oder nach dem MOS-Prinzip konzipiert, und im allgemeinen gilt diese Trennung sogar für die Geräte, die mit diesen Schaltkreisen gefertigt werden. Gegenwärtig wird bei großen elektronischen Anlagen wie Computern gewöhnlich die Bipolarvariante bevorzugt, während Quarzuhren, Taschenrechner oder

elektronisches Spielzeug eine Domäne der MOS-Technik bilden.

Im Falle von Bipolarschaltkreisen ist es unumgänglich, die verschiedenen Transistorbezirke voneinander zu isolieren, während bei MOS-Schaltungen unter Umständen darauf verzichtet werden kann, weil sich der Stromfluß dann im wesentlichen auf die Kanäle beschränkt. Abb. 8.3.7-1 zeigt, wie sich zwei komplementäre *npn*- bzw. *pnp*-Transistoren nebeneinander realisieren lassen. Die elektrische Trennung wird durch ständig gesperrte *pn*-Übergänge erreicht, von denen die wannenförmig gestalteten Funktionselemente allseitig umgeben sind. Zur Herstellung integrierter Schaltungen vom Bipolartyp ist im Vergleich zur Fertigung diskreter Bauelemente ein zusätzlicher Maskierungs- und Diffusionsschritt notwendig, um diese Trennung zu erzielen. Auch für das Aufbringen der metallischen Leiterstränge zwischen den einzelnen Bereichen ergeben sich zusätzliche Probleme, weil bestimmte Verbindungen naturgemäß von anderen isoliert sein müssen. Infolgedessen sind für die Herstellung des Netzwerks meist mehrere Verfahrensschritte des Aufbringens von Metall- bzw. Isolatorschichten unumgänglich.

Schließlich hat der Übergang zur Mikroelektronik auch ein Umdenken beim *Schaltungsentwurf* mit sich gebracht. Während man früher bemüht war, mit möglichst wenigen aktiven Bauelementen auszukommen, mit passiven Elementen jedoch nicht sparte, haben sich die Verhältnisse in der Mikroelektronik umgekehrt. Transistor- und Diodenfunktionen lassen sich einfach und billig auf den Schaltkreisen realisieren, dagegen bereitet die Fertigung genauer Widerstände Schwierigkeiten, und übliche Kapazitäten oder Induktivitäten können gar nicht eingebaut werden. Man versucht daher, weitgehend mit aktiven Elementen und wenigen Widerständen auszukommen. Bei der *Dioden-Transistor-Logik* (DTL) setzt man Transistoren und Dioden als Funktionselemente ein, während man bei der neuerdings bevorzugten *Transistor-Transistor-Logik* (TTL oder I²L) auch auf Dioden verzichtet.

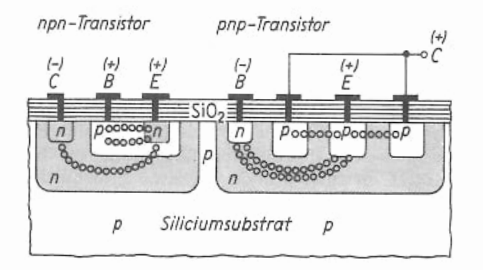


Abb. 8.3.7-1 Nebeneinanderliegende komplementäre Transistoren in einem Festkörperschaltkreis

Das Haupteinsatzgebiet der Mikroelektronik bilden digitale Logikschaltungen, mit denen sich die verschiedensten Rechen- und Speicheroperationen auf der Grundlage des binären Zahlensystems ausführen lassen. Die Anzahl der hierfür benötigten Funktionselemente ist allerdings außerordentlich groß. Infolgedessen entstand der Trend zu immer höher integrierten Schaltkreisen. Als *Integrationsgrad* definiert man gewöhnlich die Anzahl aktiver Bauelementefunktionen, die ein Festkörperschaltkreis umfaßt. Danach ist diskreten Bauelementen der Integrationsgrad 1 zuzuordnen. Die ersten Festkörperschaltkreise besaßen einen Integrationsgrad von etwa $2^6 = 64$ aktiven Bauelementen je Chip. Wenige Jahre später wurden mittelintegrierte Schaltkreise mit einem Integrationsgrad bis zu $2^{10} = 1024$ gefertigt und vor allem als Rechen- oder Zählschaltungen eingesetzt. Inzwischen sind für komplexere Rechen- und Speicherschaltungen Schaltkreise mit Integrationsgraden von 2^{12} bis 2^{14} entwickelt worden. Man spricht in diesem Zusammenhang von VLSI-Schaltkreisen [*very large scale integration* engl., sehr hoher Integrationsgrad]. In Abb. 8.3.7-2 ist die Entwicklung der Packungsdichte elektronischer Bauelemente in Geräten – ausgedrückt in aktiven Funktionselementen je Volumeneinheit – in der zeitlichen Entwicklung der Elektronik schematisch dargestellt. Aufgrund der schon jetzt abzusehenden Nachfolgeentwicklungen zur jetzigen Mikroelektronik kann erwartet werden, daß die prinzipielle Grenze noch keineswegs erreicht ist. Das beweist übrigens ein Vergleich mit der vermuteten Funktionsdichte des menschlichen Gehirns, das freilich nach gänzlich andersartigen molekularen Prinzipien arbeitet. Eine wesentliche Überlegenheit der elektronischen im Vergleich zu molekularbiologischen Logikoperationen besteht aber zweifellos in der um Größenordnungen höheren Schaltgeschwindigkeit.

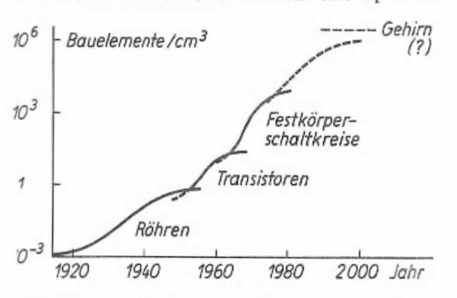


Abb. 8.3.7-2 Zeitliche Entwicklung der Packungsdichte elektronischer Bauelemente

8.3.7.2. Technologie der Mikroelektronik. Das entscheidende Glied für die Produktion mikroelektronischer Schaltkreise bildet die Entwicklung und Beherrschung einer neuartigen Technologie, die hinsichtlich Reinheit und Genauigkeit keine Parallelen in anderen Zweigen der Industrie besitzt. Eine hochproduktive Fertigung wird ermöglicht, indem bei den einzelnen Prozeßschritten die Halbleiterchips für eine größere Anzahl von Schaltkreisen gleichzeitig bearbeitet werden. Die Produktion erfolgt in klimatisierten und staubfreien Räumen. Darüber hinaus laufen die wichtigsten Prozesse innerhalb geschlossener Anlagen in einer definierten gasförmigen Umgebung bzw. im Vakuum ab. Ausgeklügelte Kontroll- und Regeleinrichtungen sorgen für die exakte Einhaltung der Verfahrensparameter, und zur Überwachung der Qualität werden hochempfindliche Testgeräte eingesetzt. Ohne auf Einzelheiten einzugehen, sollen die wichtigsten Fertigungsschritte bei der Herstellung von Festkörperschaltkreisen umrissen werden:

1. Zunächst werden die von Spezialbetrieben gelieferten Siliciumeinkristalle mittels diamantbesetzter Stahlsägen in Scheiben von 250 bis 500 µm Dicke zerteilt. Aus einem Einkristall erhält man einige hundert solcher Scheiben mit 75 bis 150 mm Durchmesser. Jede dieser Scheiben liefert die Chips für größenordnungsmäßig 100 Schaltkreise. Die Aufteilung erfolgt jedoch erst, wenn die Strukturen in und auf dem Halbleiter fertiggestellt sind.

2. In speziellen Abteilungen werden ausgehend von Zeichnungen mit Hilfe präziser optischer Abbildungs- und Verkleinerungsverfahren Schwarz-Weiß-Fotomasken auf Glasplatten hergestellt. Diese Masken stellen Negativbilder der Diffusionsfenster dar, durch die eine Umdotierung des Halbleiters vorgenommen wird. Für jeden Dotierungsschritt benötigt man gesonderte Masken. Da jede Halbleiterscheibe etwa 100 spätere Chips umfaßt und in jedes Chip je nach Integrationsgrad zwischen 10^2 und 10^5 dotierte Bereiche einzuarbeiten sind, werden an die Präzision höchste Anforderungen gestellt. Das gilt in gleichem Maße für die Wechselmasken, d. h. dünne Metallbleche mit einer Vielzahl definierter Öffnungen, die für das selektive Aufdampfen oder Aufstäuben von Metallkontakten und Leiterbahnen benötigt werden. Während die Fotomasken meist in vergrößertem Maßstab angefertigt und mittels optischer Verkleinerung auf die Halbleiterscheibe abgebildet werden, müssen die Wechselmasken die gewünschte Struktur in der Originalgröße enthalten. Bei gegenwärtig schon üblichen Rastergenauigkeiten von etwa 2 µm und einem Trend zur Beherrschung von Dimensionen im Submikrometerbereich ergeben sich bezüglich der Präzision extreme Forderungen. In zunehmendem Maße werden Wechselmasken daher durch Einsatz der Elektronenstrahllithographie hergestellt. Man bildet die Vorlage dabei mit einem rasterförmig geführten, fein fokussierten Elektronenstrahl auf das Blech ab, das mit einem lichtempfindlichen Polymerlack, auch Fotoresist genannt, überzogen ist. An den exponierten Stellen härtet dieser Lack aus. Durch Auflösen des nichtgehärteten Lacks und nachfolgendes Ätzen arbeitet man die Fenster in die Maske ein.

3. Das Einarbeiten der dotierten Bereiche geschieht nun mit dem als Fotolithographie bezeichneten Verfahren, das charakteristisch für die Mikroelektronik ist. Die Siliciumscheiben werden zunächst gereinigt und bei Temperaturen um 1200 °C in Quarzrohröfen durch Einwirkung von Sauerstoff oberflächlich oxydiert. Auf die Scheiben wird einseitig Fotolack durch Schleudern gleichmäßig aufgetragen. In speziellen Projektionseinrichtungen erfolgt nun die Abbildung der Fotoschablonen bei optischer Verkleinerung. Äußerst wichtig ist dabei eine genaue Positionierung, damit die relative Lage der einzelnen Bereiche im Verlauf der nachfolgenden Prozeßschritte genau eingehalten werden kann. Nach dem Entwickeln und Auflösen des nichtexponierten Lacks verbleibt die Lackschicht nur an den belichteten Stellen.

4. Im weiteren Fortgang wird durch selektive Ätzung die Siliciumoxidschicht an den nicht belichteten Stellen abgetragen und auch der ausgehärtete Lack entfernt. Die früher übliche naßchemische Behandlung mit geeigneten Säuremischungen wird zunehmend durch das genauere *Plasmaätzen* verdrängt. Man setzt die Scheiben dabei einer elektrischen Plasmaentladung aus, die bei niederem Druck in Dämpfen von Fluorkohlenwasserstoffen wie CH_3F aufrechterhalten wird. Der Abtrag von SiO_2 erfolgt hauptsächlich durch die chemische Wirkung der auftreffenden Molekülfragmente und -Ionen. Zur Entfernung des gehärteten Lacks eignet sich eine Plasmaätzung in Sauerstoff unter vermindertem Druck.

5. Die so vorbereiteten Scheiben werden in Rohröfen der Einwirkung von Dämpfen ausgesetzt, in denen die Dotierungselemente, z. B. Bor oder Phosphor, enthalten sind. Die Tiefe und das Konzentrationsprofil der Eindiffusion von Dotanten lassen sich durch Regelung des Temperaturverlaufs sehr genau kontrollieren. Beispielsweise diffundieren Phosphoratome bei 1100 °C in einer Stunde um etwa 1 µm in das Siliciumgitter ein. Gegenwärtig werden Toleranzen von wenigen Zehntel µm in der Ausdehnung der dotierten Bereiche beherrscht.

Im allgemeinen müssen die unter 3. bis 5. aufgeführten Prozeßschritte bei der Fertigung von Festkörperschaltkreisen mehrfach unter leicht abgewandelten Bedingungen wiederholt werden,

bevor die Halbleiterstruktur fertiggestellt ist. Besonders für hochwertige Bipolarschaltkreise arbeitet man häufig mit der Variante der Epitaxietechnik (↑ 8.3.6.2.2.). In diesem Falle läßt man auf die Siliciumscheiben zunächst durch Behandlung mit $SiCl_4$- oder $SiHCl_3$-haltigen Gasmischungen bei hohen Temperaturen eine epitaktische Si-Schicht von einigen 10^1 µm Dicke aufwachsen. Die dotierten Bereiche werden wie oben beschrieben in diese Schicht eingearbeitet. Die ursprünglichen Halbleiterscheiben dienen dann letztlich nur noch als mechanisch stabile Träger der Schaltung.

6. Nach dem letzten Dotierungsschritt werden die Si-Scheiben erneut oberflächlich oxydiert. Im Ergebnis eines weiteren fotolithographischen Prozesses werden die Teile der Halbleiteroberfläche freigelegt, an denen metallische Kontakte anzubringen sind. Unter Benutzung der in 2. erwähnten Wechselmasken erfolgt das Anbringen der meist aus Aluminium und Gold bestehenden Kontakte durch Aufdampfen oder Aufstäuben im Hochvakuum.

7. Ebenfalls durch Aufdampfen oder Aufstäuben werden die Leiterbahnen mit Hilfe weiterer Wechselmasken abgeschieden. Um Bahnen voneinander isolieren zu können, sind hierbei häufig mehrere Schritte erforderlich, wobei zwischendurch eine selektive Abscheidung einer isolierenden Schicht (SiO_2, manchmal auch Si_3N_4) vorgenommen wird.

8. Erst jetzt erfolgt das Aufteilen der Scheiben in die Chips der einzelnen Schaltkreise. Diese Chips werden in Plast- oder Keramikgehäuse montiert, und durch Anschweißen oder „Anbonden" feiner Golddrähte mit Spezialmaschinen wird der Anschluß an die Sockelstifte vorgenommen. Schließlich verkapselt man die fertigen Schaltkreise luftdicht unter Benutzung von Polymerharz.

8.3.7.3. Trends der Mikroelektronik. Bei der festkörperphysikalischen und mikroelektronischen Forschung laufen intensive Bemühungen, die Miniaturisierung und die Schaltgeschwindigkeit durch Verbesserung der jetzigen Schaltkreise oder durch Nutzung neuer Prinzipien weiter voranzutreiben. Auf einige erkennbare Trends soll hingewiesen werden:

Zweifellos lassen sich Schaltkreise mit höherer Packungsdichte auf der Grundlage der bekannten Bipolar- oder MOS-Technik mit Silicium als Werkstoff realisieren, wenn es gelingt, die lithographische Mikrostrukturierung weiter zu verbessern und das Rastermaß von gegenwärtig etwa 2 µm auf einige Zehntel µm herabzusetzen. Neue Verbesserungen der Fotolithographie, beispielsweise das Aufbringen organischer Zwischenschichten oder der Einsatz anorganischer Fotoresists wie Ag/Ge/Se anstelle der bisherigen Fotolacke versprechen eine Strukturierungsgenauigkeit von 0,3 bis 0,6 µm. Bei noch kleineren Dimensionen kann sichtbares Licht wegen der durch die Wellenlänge begrenzten Abbildungsgenauigkeit nicht mehr eingesetzt werden. Dagegen könnte man erheblich größere Genauigkeiten durch Verwendung weicher Röntgenstrahlen oder Elektronenstrahlen erreichen. Mit diesem Ziel wird an der Entwicklung einer *Röntgen-* bzw. *Elektronenstrahllithographie* gearbeitet. Auch Versuche zum direkten „Einschreiben" der Dotierung mittels fokussierter *Ionenstrahlen* sind in diesem Zusammenhang anzuführen.

Bei Dimensionen um oder unter 1 µm wirken sich kleine Unebenheiten an den Kontakten störend aus, die durch Legierungsbildung zwischen Silicium und Aluminium hervorgerufen werden. Zwischenschichten leitfähiger Silicide, beispielsweise $PdSi_2$ oder $CrSi_2$, lassen in naher Zukunft wesentliche Verbesserungen erwarten.

In weiterer Zukunft kann der Übergang zu Schaltkreisen auf der Basis von $A^{III}B^{V}$-Verbindungen bezüglich der Schaltgeschwindigkeit wesentliche Fortschritte bringen, da die Beweglichkeit der Elektronen in diesen Materialien erheblich größer ist als in Si. Diese Überlegenheit könnte noch verbessert werden, wenn man die Schaltungen bei tiefen Temperaturen betreibt (s. a. 8.3.6.4.4.).

Die bisherigen Festkörperschaltkreise werden entsprechend der Planartechnik als Schichtsysteme in zwei Dimensionen ausgeführt. Eine Erweiterung zum mehrlagigen Schaltkreis würde das Aufbringen einer weiteren Epitaxieschicht auf der abschließenden amorphen Isolatorschicht erforderlich machen. Eine mögliche Lösung dieses Problems zeigt die als *Graphoepitaxie* in den letzten Jahren bekanntgewordene Technik auf. Danach ist eine Abscheidung einkristalliner Schichten auf amorphen Substraten wie SiO_2 möglich, wenn man in die Oberfläche des Substrats vorher mit den Mitteln der Lithographie ein feines Rechteckrelief paralleler Gräben einarbeitet. Dieser interessante Effekt, der mit einer bevorzugten Orientierung primärer Kristallkeime an den unteren Grabenkanten zusammenhängt, dürfte auch für Kombinationen zwischen Halbleiterbauelementen und optischen, magnetischen oder ferroelektrischen Funktionselementen Interesse erlangen.

Eine Alternative zur Halbleitertechnik entwickelt sich gegenwärtig mit der *Mikrokryoelektronik*. Hierbei werden die Logikoperationen durch das Schalten supraleitender Metallschichtelemente auf der Basis des Josephsoneffekts (↑ 8.5.4.) ausgeführt. Vorteile ergeben sich aus der hohen Schaltgeschwindigkeit im Pikosekundenbereich und aus den kleinen Ansteuerspannungen (< 100 mV). Da die Supraleitung jedoch nur bei sehr tiefen Tempe-

raturen auftritt, müßte ein extrem schneller und hochintegrierter „Supracomputer“ ständig auf der Temperatur des flüssigen Heliums gehalten werden. Immerhin erscheinen Konzeptionen unter Benutzung geschlossener Kreislaufkühlaggregate realisierbar. Die endgültige Begrenzung der Rechengeschwindigkeit eines solchen Computers würde wahrscheinlich durch die von der Lichtgeschwindigkeit bestimmte Signalausbreitung bestimmt!

Abschließend bleibt festzustellen, daß die Mikroelektronik die bisher fortgeschrittenste und konsequenteste Anwendung der Erkenntnisse vom atomistischen Aufbau der Materie verkörpert.

8.4. Leiter

8.4.1. Kristallstruktur und metallische Bindung

8.4.1.1. Kennzeichen eines Metalls. Mehr als drei Viertel aller Elemente liegen bei Zimmertemperatur im metallischen Zustand vor. Das ist natürlich kein Zufall, sondern eng mit dem Aufbau der Elektronenhülle der Atome selbst verknüpft (↑ 4.2.3.). Metalle und metallische Legierungen zeichnen sich durch bestimmte Eigenschaften aus, die für ihren Werkstoffeinsatz einzeln oder im Komplex entscheidend sind. Hervorzuheben sind inbesondere die hohe Konzentration beweglicher Elektronen, mit der die hohe elektrische Leitfähigkeit, das gute Reflexionsvermögen und die hohe Wärmeleitfähigkeit sowie die plastische Verformbarkeit verbunden sind. Analysiert man das Verhalten von Metallen näher, dann können im allgemeinen folgende Feststellungen getroffen werden: Metalle besitzen spezifische Widerstände zwischen 1 und 100 $\mu\Omega$cm; der Widerstand nimmt mit wachsender Temperatur zu; der Zusammenhang zwischen der Stromdichte und der elektrischen Feldstärke ist linear (Ohmsches Gesetz); bei tiefen Temperaturen besitzen Metalle einen endlichen Restwiderstand (Mathiessensche Regel); bei sehr tiefen Temperaturen werden viele Metalle supraleitend; unter der Einwirkung von elektrischen und magnetischen Feldern sowie von Temperaturgradienten treten thermogalvano-magnetische Effekte auf; Metalle absorbieren und reflektieren Licht sehr stark; sie kristallisieren in Gittern hoher Symmetrie und hoher Packungsdichte; bei geeigneten Temperaturen lassen sie sich plastisch verformen; die mechanischen Eigenschaften umfassen einen breiten Wertebereich.

Das macht auch die an sich unerschöpfliche Vielfalt metallischer Konstruktionswerkstoffe aus und eröffnet viele Möglichkeiten der Anwendung als Leiter, Supraleiter und als Reflektormaterial.

8.4.1.2. Kristallstruktur der Metalle. Bedingt durch die Eigenschaften der metallischen Bindung (s. a. 5.1.1.2.1.) – nicht absättigbar, isotrop, große Reichweite – werden von Metallen Gitter bevorzugt, die eine hohe Koordinationszahl aufweisen. Das sind für drei Viertel aller metallischen Elemente das kubisch raumzentrierte (*krz*), das kubisch flächenzentrierte (*kfz*) und das hexagonale Gitter (s. a. 6.4.1.). Geradezu typisch ist das Auftreten der dichtesten Kugelpackung, da diese der metallischen Bindung am angemessensten ist. Unter Normalbedingungen bilden ein *krz-Gitter* die metallischen Elemente: Li, Na, K, Rb, Cs, Fr, Ba, Ra, V, Nb, Ta, Cr, Mo, W, Fe und Eu, ein *kfz-Gitter*: Sr, Ca, Rh, Ir, Ni, Pd, Pt, Cu, Ag, Au, Al, Pb, Ac, Ce, Th, Am, Yb und ein *hexagonales Gitter* besitzen, geordnet nach zunehmender Abweichung von der hexagonal dichtesten Kugelpackung: Re, Tc, Co, Mg, Sc, Y, Ti, Zr, Hf, Ru, Os, Be, Gd, Tb, Dy, Ho, Er, Tm, Lu, Tl, Zn, Cd, La, Pr, Nd, Cm und Sm. Einige andere Metalle kristallisieren in Gittern niedriger Symmetrie. Diese besitzen dann auch als Einkristall eine bestimmte Anisotropie der physikalischen Eigenschaften, z. B. der Leitfähigkeit, bezüglich verschiedener Kristallorientierung zum elektrischen Feld. Durch Variation der Zustandsgrößen Druck und Temperatur kann oft eine Änderung der Kristallstruktur eines metallischen Elements erzwungen werden, d. h. oft existieren in anderen Zustandsbereichen andere Gleichgewichtsbedingungen innerhalb der Nah- und Fernordnung (s. a. 6.4.) der Bausteine. Man bezeichnet diese Erscheinung als *Polymorphismus*. Aus der Vielfalt der Beispiele sei der Modifikationswechsel des Zinns, α-Sn (Diamantgitter)

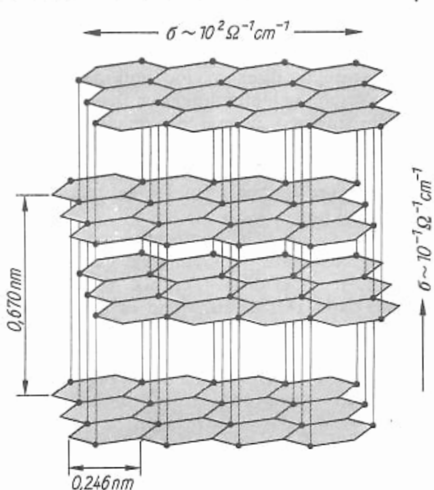


Abb. 8.4.1-1 Struktur des Graphits

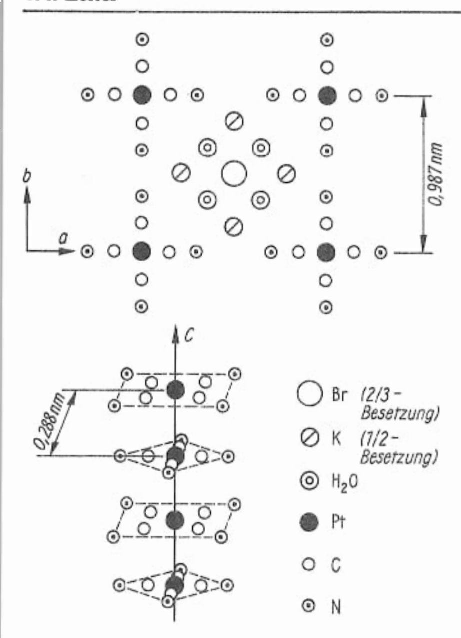


Abb. 8.4.1-2 Relative Lage von Molekülen im KCP nach MINOT und PERLSTEIN

$\xrightleftharpoons{13{,}2\,°C}$ β-Sn (hexagonal), angeführt, bei dem gleichzeitig ein starker Umschlag von Halbleiterverhalten zu metallischem Verhalten beobachtet wird. An diesem Beispiel wird auch deutlich, daß die Zuordnung zu einem bestimmten Leitungstyp konkret für einen definierten Zustand erfolgen sollte. Man kann z. B. auch prinzipiell erwarten, daß sich nahezu alle Elemente unter extrem hohem Druck metallisch verhalten. Paradebeispiel dafür ist der metallische Wasserstoff. Oft beobachtet man auch eine ganze Reihe reversibler polymorpher Übergänge, wie beim Cs:
Cs I (*krz*) $\xrightleftharpoons{23\,GPa}$ Cs II (*kfz*) $\xrightleftharpoons{42\,GPa}$ Cs III (*kfz*) $\xrightleftharpoons{125\,GPa}$ Cs IV. Je lockerer ein Metallkristall zunächst gepackt ist, desto wahrscheinlicher liegt der erste Phasenübergang bei einem relativ niedrigen Druck. Auch aus den Halbleitern Germanium und Silicium, den $A^{III}B^{V}$- bzw. $A^{II}B^{VI}$-Verbindungshalbleitern entsteht unter hohem Druck eine metallische Phase. Umgekehrt werden einige Metalle wie Ytterbium, Calcium und Strontium bei Druckerhöhung zunächst halbleitend. Extreme Drücke führen auch zu einer Umordnung der Elektronenhülle.

8.4.1.3. Metallische Bindung. Die metallische Bindung zeichnet sich dadurch aus, daß im Gitter eines metallischen Elements oder einer Metallegierung Elektronen auftreten, die nicht mehr an einen bestimmten Atomrumpf gebunden sind. In den meisten Fällen gibt jeder Gitterbaustein ein bis zwei Elektronen an das Kristallgitter ab. Diese Valenz- oder *Leitungselektronen* bilden ein *Elektronengas*, das sich relativ leicht im Kristall bewegen kann. Die Existenz dieser nahezu freien Elektronen läßt sich direkt experimentell nachweisen. Man nutzt dazu im *Tolmanversuch* ihre Trägheit aus, die sich bei der Beschleunigung eines Metallstücks bemerkbar macht. Die Bindungskraft entsteht durch kollektive Wechselwirkung der positiven Atomrümpfe und quasifreien Elektronen. Es muß besonders hervorgehoben werden, daß für die metallischen Eigenschaften entscheidend die Nahordnung verantwortlich ist. Also zeigen auch mikrokristalline und amorphe Metalle diese Eigenschaften, und es gibt bei höheren Temperaturen viele flüssige Metalle.

8.4.1.4. Lineare Metalle. Es überrascht niemand, wenn man Graphit mit seiner ausgeprägten Schichtstruktur (Abb. 8.4.1-1) zu den Metallen zählt. Hier wird die zweidimensional ausgebildete hohe Leitfähigkeit einer konjugierten Molekülstruktur als Kriterium herangezogen, obwohl strenggenommen kein metallischer Zustand vorliegt. Die Bindungskräfte im Graphitgitter sind in der Schichtebene kovalenten Typs und zwischen den molekularen Schichten

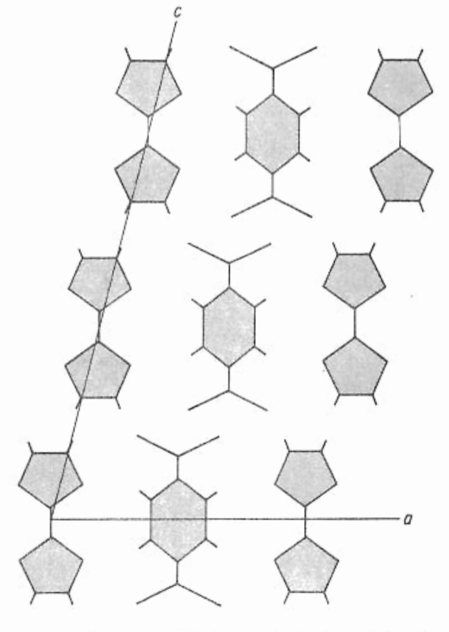


Abb. 8.4.1-3 Lage der Moleküle im 1 : 1-Ladungsübertragungskomplex des TTF-TCNQ

vom van-der-Waals-Typ. Ganz anders sieht das bei einigen anorganischen und organischen linearen Strukturen aus, die tatsächlich in einer bevorzugten Kristallrichtung alle Besonderheiten des metallischen Zustands zeigen. Leicht einzusehen ist das beispielsweise beim $K_2Pt(CN)_4Br_{0,3} \cdot 2{,}3\,H_2O$ oder kurz *KCP*. Charakteristisch für diese Stoffe ist die Stapelung planarer Moleküle, wobei eine lineare Metallkette entsteht (Abb. 8.4.1-2). Der *lineare metallische Zustand* läßt sich dann realisieren, wenn es gelingt, eine lineare Anordnung von irgendwie gearteten Gitterbausteinen zu schaffen, die einen solchen Elektronenüberschuß oder auch Elektronenmangel besitzen, daß sich ein nicht vollbesetztes eindimensionales Leitungsband ausbilden kann. Das wurde seit 1973 z. B. durch die Synthese von Ladungsübertragungskomplexen des Typs Tetrathiafulvalen – Tetracyanochinodimethan, kurz *TTF–TCNQ*, erreicht (Abb. 8.4.1-3). Hier sind überhaupt keine metallischen Elemente beteiligt, sondern es liegen gleichnamige molekulare Ionenketten vor.

8.4.2. Leitereigenschaften von Metallen

8.4.2.1. Ohmsches und Wiedemann-Franzsches Gesetz. 1826 fand Ohm die Proportionalität zwischen Stromstärke und angelegter Spannung bei Metallen. Nach dem *Ohmschen Gesetz* gilt für den Zusammenhang zwischen der Stromdichte j und der elektrischen Feldstärke E die lineare Beziehung

(1) $j = \sigma E.$

Die *spezifische Leitfähigkeit* berechnet sich für einen zylindrischen metallischen Widerstand R der Länge l und der Querschnittfläche A zu $\sigma = 1/\varrho = l/(RA)$.
Der Kehrwert der spezifischen Leitfähigkeit σ wird *spezifischer Widerstand* ϱ genannt und ist eine wichtige Materialgröße. Für einige wenige

Tab. 8.4.2-1 Spezifischer Widerstand ϱ in $\mu\Omega$cm und Temperaturkoeffizient α in K^{-1} einiger wichtiger Leiter

Leiter	ϱ	$\alpha \cdot 10^3$
Silber	1,6	4
Kupfer	1,8	4
Gold	2,2	4
Aluminium	2,7	4
Molybdän	5,4	5
Wolfram	5,5	4
Zinn	10	5
Blei	21	4
Quecksilber	96	1

Anwendungsfälle ist zu beachten, daß σ Tensorcharakter besitzt. Für alle praktischen Fälle genügt es, entsprechend der isotropen Kristallorientierung in metallischen Leitern den skalaren Mittelwert des Tensors der spezifischen Leitfähigkeit einzusetzen. Die Wärmeleitfähigkeit λ und σ sind miteinander durch das zunächst empirisch gefundene und später theoretisch begründete *Wiedemann-Franzsche Gesetz*

(2) $\dfrac{\lambda}{\sigma T} = L$

verbunden. Die *Lorenzzahl* L ist für metallische Elemente nahezu eine Konstante und beträgt $L = 2{,}5 \cdot 10^{-8}\ V^2K^{-2}$.

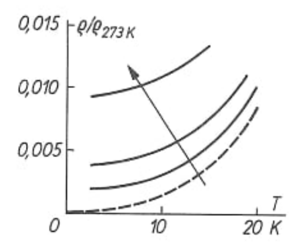


Abb. 8.4.2-2 Restwiderstand von Goldproben verschiedener Reinheit; der Verunreinigungsgehalt steigt in *Pfeilrichtung*, *gestrichelt* theoretische Kurve des Idealkristalls

8.4.2.2. Temperaturabhängigkeit des spezifischen Widerstands und der Wärmeleitfähigkeit. Für die elektrische Leitung benötigt man in den meisten Fällen eine hohe Leitfähigkeit. Bei vielen Anwendungen metallischer Leiter ist es auch wichtig, die Temperaturabhängigkeit des spezifischen Widerstands $\varrho(T)$ zu nennen. Als gut handhabbare Materialgröße definiert man den Temperaturkoeffizienten von ϱ als

(3) $\alpha = \dfrac{1}{\varrho}\dfrac{d\varrho}{dT}.$

Bei Metallen ist α positiv, d. h. mit steigender Temperatur nimmt ϱ zu (Tab. 8.4.2-1).
Nach der *Matthiessenschen Regel*

(4) $\varrho = \varrho(T) + \varrho_0$

lassen sich zwei Anteile des spezifischen Widerstands unterscheiden. Der Restwiderstand ϱ_0 ist temperaturunabhängig und bestimmt diese Materialgröße bei sehr tiefen Temperaturen im normalleitenden Zustand (Abb. 8.4.2-2). Er geht auf die Wirkung von Gitterstörungen wie Verunreinigungen, Leerstellen oder Versetzungen zurück. Aus der Messung des spezifischen Widerstands bei sehr tiefen Temperaturen kann man wertvolle Aussagen über das Aus-

maß der Gitterstörungen und über die Reinheit von Metallen gewinnen. Der temperaturabhängige Anteil $\varrho(T)$ wird dagegen durch die Streuung von Elektronen an den Gitterschwingungen verursacht. Deshalb würde für einen idealen Kristall $\varrho = \varrho(T)$ gelten.

Analog zum spezifischen Widerstand läßt sich der *spezifische Wärmewiderstand* $w = 1/\lambda$ als Kehrwert der Wärmeleitfähigkeit einführen. Sein Gesamtwert $w = w_{el} + w_G$ setzt sich additiv aus dem Elektronenanteil w_{el} und dem Gitteranteil w_G zusammen. Der Elektronenanteil $w_{el} = aT^2 + \varrho_0/(LT)$ wird durch die gleichen Streuvorgänge bestimmt wie der elektrische Widerstand. Bildet man das thermische Widerstandsprodukt $\zeta = Tw_{el}$, dann kann nach BORELIUS eine zu 8.4.(4) analoge Beziehung

(5) $\zeta = \zeta(T) + \zeta_0$

aufgestellt werden. Dabei ist $\zeta(T)/T$ der temperaturabhängige Anteil des elektronischen Wärmewiderstands des idealen Kristalls und ζ_0 ein temperaturunabhängiger elektronischer Anteil. Für gute Leiter gilt $\lambda_{el} \gg \lambda_G$, d. h., fast die gesamte Wärmemenge wird durch Leitungselektronen transportiert. Der Gitteranteil der Wärmeleitfähigkeit wird dagegen durch Gitterschwingungen oder Phononen bewirkt. Diese Quasiteilchen können die verschiedensten Streuvorgänge erleiden. So findet man z. B. die folgenden Temperaturabhängigkeiten für die Streuung von Phononen: an Punktdefekten: $\lambda_G \sim 1/T$, an Versetzungen: $\lambda_G \sim T^2$, an Kristallgrenzen: $\lambda_G \sim T^3$ und an Leitungselektronen: $\lambda_G \sim T \ldots \lambda_G \sim T^2$.

8.4.2.3. Thermoelektrische, galvanomagnetische und thermomagnetische Effekte. Durch einen Temperaturgradienten wird in einem Leiter die räumliche Verteilung der Elektronen verändert, d. h. in geringem Maße wird die Konzentration der Leitungselektronen ortsabhängig. Infolge eines solchen Konzentrationsgefälles kommt es zur Ausbildung einer Potentialdifferenz. Diese Erscheinung wird als *Seebeckeffekt* bezeichnet und gehört zusammen mit dem *Peltiereffekt* zu den *thermoelektrischen Effekten.* Der Seebeckeffekt wird vor allem mit Hilfe von Thermoelementen zur Bestimmung von Temperaturdifferenzen verwendet. Beim Peltiereffekt wird dagegen durch Anlegen einer Potentialdifferenz an einen Leiter ein Temperaturgradient erzeugt. Auf der Basis dieses Effekts arbeitende Anordnungen können für die *thermoelektrische Kühlung* eingesetzt werden. Besonders günstige Kühleffekte um 50 bis 80 K erzielt man mit Kombinationen der Metalle Bismut oder Antimon mit den Halbleitern Selen oder Tellur (↑ 8.3.5.2.).

Wirkt ein magnetisches Feld auf einen Leiter ein, in dem Elektronen fließen, dann erfolgt durch die Lorentzkraft (↑ 2.2.3.) eine Bahnablenkung der Leitungselektronen. Man kann also wie beim *Halleffekt* (↑ 8.3.5.2.) senkrecht zum primären elektrischen Feldvektor und zum Vektor der magnetischen Induktion eine sekundäre Potentialdifferenz beobachten. Der Halleffekt von Metallen wird insbesondere zur Messung starker Magnetfelder verwendet. Die Bahnkrümmung der Leitungselektronen führt auch zur Erhöhung des spezifischen Widerstands eines Leiters im Magnetfeld. Diese Erscheinung wird als Magnetowiderstandseffekt bezeichnet. Sie ist beim Bismut so groß, daß Widerstände aus Bismut zur Messung von Magnetfeldern benutzt werden können. Eine kombinierte Einwirkung von elektrischen und magnetischen Feldern sowie von Temperaturgradienten führt unter Berücksichtigung der speziellen Winkelbeziehungen zwischen Feldern und Temperaturgradienten bei Metallen zu einer ganzen Reihe *thermo-galvano-magnetischer* Effekte (Tab. 8.4.2-3).

Tab. 8.4.2-3 Galvanomagnetische, thermoelektrische und thermogalvanomagnetische Effekte; Indizes x, y und z geben Richtung des elektrischen Felds $\boldsymbol{E}$ der magnetischen Induktion $\boldsymbol{B}$ und der Temperaturdifferenz ΔT in einem rechtwinkligen Koordinatensystem an; $\Delta\varrho$ und $\Delta\lambda$ Änderung des spezifischen Widerstands bzw. der Wärmeleitfähigkeit

Einwirkung auf den Leiter	Ergebnis	Name
ΔT_x	E_x	Seebeckeffekt
E_x	ΔT_x	Peltiereffekt
E_x, B_y	E_z	Halleffekt
E_x, B_y	$\Delta\varrho_x$	Magnetowiderstandseffekt
E_x, B_y	ΔT_z	Ettingshauseneffekt
E_x, B_y	ΔT_x	Nernsteffekt
ΔT_x, B_y	ΔT_z	Righi-Leduc-Effekt
ΔT_x, B_y	$\Delta\lambda_x$	Maggi-Righi-Leduc-Effekt
ΔT_x, B_y	E_z	1. Ettingshausen-Nernst-Effekt
ΔT_x, B_y	E_x	2. Ettingshausen-Nernst-Effekt

8.4.2.4. Spezifischer Widerstand von Legierungen. *Metallische Legierungen* können in verschiedenen Formen auftreten. Bei binären Legierungen wird man entsprechend den Grundtypen der Zustandsdiagramme (↑ 6.4.5.) ein sehr unterschiedliches Verhalten bezüglich elektronischer Leitungsvorgänge erwarten. *Kristallgemische*, bei denen die Legierungskomponenten nur sehr begrenzt miteinander mischbar

sind, wirken elektrisch ähnlich wie eine dreidimensionale stochastische Anordnung aus zwei verschiedenen Widerstandstypen. In Abhängigkeit von der Konzentration der einen oder anderen Komponente wird ein Verhalten wie in Abb. 8.4.2-4 links typisch sein, das nach LICHTENECKER durch

$$\varrho = \varrho_A^c \varrho_B^{1-c} \tag{6}$$

beschrieben wird.

bei entsprechenden Konzentrationswerten, die den Mischungsverhältnissen $N_A : N_B = m : n$ bzw. $q : p$ usw. entsprechen, Minima der ϱ-Werte (Abb. 8.4.2-4 rechts). Insbesondere werden mischkristalline Legierungen verwendet, um Widerstandswerkstoffe zu gewinnen, die einen verschwindend kleinen Temperaturkoeffizienten α besitzen (Tab. 8.4.2-5). Der physikalische Ausgangspunkt sind die Beziehungen 8.4.(3) und 8.4.(4). Man sieht sofort ein, daß man

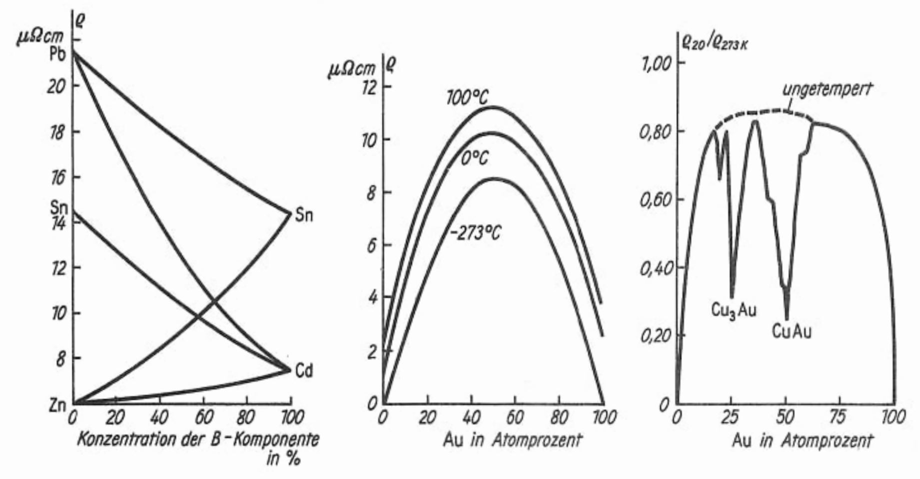


Abb. 8.4.2-4 Spezifischer Widerstand ϱ binärer Legierungen; *links* Kristallgemisch, *Mitte* Mischkristall Ag/Au, *rechts* intermetallische Phasen der Cu/Au-Legierung nach POSPISIL; *gestrichelte Kurve* ungetempert, *durchgezogene Kurve* getempert

Bei *Mischkristallen* (↑ 6.4.4.) bewirkt die gelöste Komponente B oder A in der Hauptkomponente A bzw. B auf jeden Fall eine zusätzliche Streuung der Leitungselektronen und damit eine Erhöhung des Gitteranteils des spezifischen Widerstands ϱ_0. Ein maximal gestörtes Gitter wird etwa bei $c = 0{,}5$ vorhanden sein, $c = N_B/(N_A + N_B)$ ist die Konzentration der Komponente B. N_A bzw. N_B ist die Zahl der Atome jeder Komponente im Mischkristall (Abb. 8.4.2-4 mitte). Bilden zwei oder auch mehr Komponenten A, B, C usw. eine *intermetallische Phase* (↑ 6.4.4.) A_mB_n, $AqBp$ usw., m, n, q und p sind ganze Zahlen, so erhält man durch starke Erhöhung des Restwiderstands ϱ_0 den Temperaturkoeffizienten α wesentlich erniedrigen kann.
Denn es gilt

$$\alpha = \frac{1}{\varrho(T) + \varrho_0} \frac{\mathrm{d}\varrho(T)}{\mathrm{d}T}. \tag{7}$$

Besonders viele Gitterstörungen erhält man bei mehrkomponentigen Systemen. Sehr variable spezifische Widerstandswerte ergeben sich auch durch Kristallgemische eines Metalls mit einem Isolator. Dadurch, daß sich die beiden Komponenten sehr stark im Widerstand unterscheiden, überstreicht das Mischsystem für ϱ viele Grö-

Tab. 8.4.2-5 Spezifischer Widerstand ϱ in μΩcm und Temperaturkoeffizient α in K^{-1} von Widerstandslegierungen

Legierung	Zusammensetzung	ϱ	$\alpha \cdot 10^5$
Chromnickel	0,8 Ni, 0,2 Cr	110	20
Neusilber	0,6 Cu, 0,17 Ni, 0,23 Zn	30	20
Konstantan	0,54 Cu, 0,45 Ni, 0,12 Mn	50	3
Manganin	0,8 Cu, 0,02 Ni, 0,12 Mn	43	2

ßenordnungen. Derartige Stoffe werden *Cermets* genannt (ceramic metal) und bestehen z. B. aus Chrom und Siliciumoxid.

8.4.3. Elektronensystem in Metallen

8.4.3.1. Klassische Elektronentheorie. Faßt man die Leitungselektronen eines Metalls als Bestandteil eines Elektronengases auf, das zunächst keinen weiteren Bedingungen unterworfen ist, so läßt sich der Gleichverteilungssatz der *klassischen Maxwell-Boltzmann-Statistik* auf dieses Elektronensystem anwenden. Auf jeden Freiheitsgrad der Translation entfällt danach ein Energieanteil $E = k_B T/2$, wobei k_B die Boltzmannkonstante ist. Das heißt, die mittlere kinetische Energie eines Leitungselektrons wäre somit

$$(8) \quad E = \frac{m}{2} v^2 = \frac{3}{2} k_B T$$

und nur von der Temperatur abhängig. Die gesamte kinetische Energie der Leitungselektronen müßte bei $T = 0$ K verschwinden. Mit diesen heute als grundsätzlich falsch erkannten Vorstellungen lassen sich das *Ohmsche Gesetz* und das *Wiedemann-Franzsche Gesetz* deuten, weil bei höheren Temperaturen klassische Beziehungen in diesen Ausnahmefällen eine gute Näherung darstellen.

8.4.3.2. Deutung des Ohmschen Gesetzes. Enthält ein Metall n Leitungselektronen pro cm^3, müßte man bei Anlegen einer elektrischen Feldstärke $|E|$ eine Stromdichte

$$(9) \quad j = \frac{dQ}{dt\, dA} = \frac{en\, dA\, v_D\, dt}{dt\, dA} = \frac{n\, e^2\, |E|\, t}{m}$$

erwarten. Dabei sollen die verwendeten Bezeichnungen v_D – die entgegen der Feldrichtung gerichtete Driftbewegung der Elektronen, dA – die Querschnittfläche des Leiters, m – die Elektronenmasse und dQ die driftende Ladungsmenge sein. Die Beziehung 8.4.(9) widerspricht nun jeder Erfahrung, da auch nach sehr langen Zeiten die Stromdichte endlich bleibt. Aus sehr langen Drähten treten keinesfalls schon bei kleinen Feldstärken nach langen Zeiten wegen

$$(10) \quad v_D = \frac{e\,|E|}{m}\, t$$

schnelle Elektronen aus. Die Erfahrung fordert also, daß die Leitungselektronen im Leiter Stöße erleiden, bei denen sie im Mittel nahezu ihre ganze Driftgeschwindigkeit verlieren. Auch ohne anliegendes elektrisches Feld führen die Elektronen durch derartige Stöße mit den Gitterbausteinen eine völlig regellose Zickzackbewegung aus. Im Mittel wird also ein Elektron im Leiter ein Stück l, die mittlere freie Weglänge, ungehindert zurücklegen und auch driften könnten. Für eine Zeit $\tau = l/\bar{v}$; $\bar{v} \gg v_D$, die mittlere freie Flugzeit, wird das Leitungselektron maximal beschleunigt auf $v_{D,max} = -\frac{eE\tau}{m}$.

Für die Größe v_D in 8.4.(9) muß man deshalb

$$(11) \quad v_D = -\frac{1}{\tau}\int_0^\tau \frac{eE}{m}\, t\, dt = -\frac{eE\tau}{2m}$$

einsetzen, da es zu jedem betrachteten Zeitpunkt Elektronen mit Driftgeschwindigkeiten $0 \leqq v_D \leqq v_{D,max}$ gibt. Das Ohmsche Gesetz ergibt sich dann einfach zu:

$$(12) \quad j = -\frac{n\, e^2 \tau}{2m} E.$$

Der Vergleich mit 8.4.(1) liefert sofort für die spezifische Leitfähigkeit bzw. für den spezifischen Widerstand

$$(13) \quad \sigma = 1/\varrho = n\, e^2 \tau / 2m.$$

Die Größe τ ist eine Materialgröße, die sowohl von der Temperatur als auch vom Bau des Kristallgitters abhängt. In τ verstecken sich die Abhängigkeiten, die schließlich zur Matthiessenschen Regel führen. Es ist interessant, die hier verwendeten Größen einmal für einen guten Leiter zu untersuchen. Für Kupfer ist ein spezifischer Widerstand von $\varrho = 1{,}8\ \mu\Omega$cm angegeben (↑ Tab. 8.4.2-1). Ein cm^3 Kupfer enthält $8{,}5 \cdot 10^{22}$ Atome, weil die Abmessung der Elementarzelle des kfz-Gitters, die vier Kupferatome umfaßt, $(0{,}3615\ \text{nm})^3$ beträgt. Wir rechnen hier damit, daß jedes Atom ein Elektron an das Gitter abgibt. Damit folgt für $\tau = 2m/e^2\, n\varrho = 0{,}46 \cdot 10^{-13}$ s. Einer freiliegenden Kupferleitung von 5 mm^2 Querschnitt kann man bei 20 °C maximal eine Stromdichte von 30 A mm^{-2} zumuten. Das heißt aber, es sind in guten Leitern etwa maximale elektrische Feldstärken von $|E|_{max} = j_{max}\varrho = 5{,}4$ mV cm^{-1} möglich. Für die thermische Geschwindigkeit nach dem Gleichverteilungssatz findet man $\langle v^2 \rangle^{1/2} = 3k_B T/m$. Das ergibt für $T = 300$ K eine Geschwindigkeit von 117 km s^{-1}, während man in dem gewählten Beispiel für die mittlere Driftgeschwindigkeit maximal $v_D = 2{,}2$ mm s^{-1} errechnet. Das ist insofern ein überraschendes Ergebnis, weil die Information, daß ein Einschalt- oder Ausschaltvorgang am Leiteranfang erfolgte, am Leiterende mit Lichtgeschwindigkeit ankommt, aber die Ladungsträger selbst nur eine Strecke von maximal 7,9 m pro Stunde zurücklegen. Frei fliegen kann der Ladungsträger aber nur $l = 5{,}4$ nm weit. Das entspricht im Kupfer einer Strecke von 15 Gitterabstän-

den. Pro zurückgelegtem Meter in Driftrichtung unter der Wirkung einer maximal möglichen Feldstärke wird der Ladungsträger dabei 185 Millionen Mal durch Stöße abgebremst. In diesem vereinfachten Bild findet man für die als *Beweglichkeit* $\mu = \bar{v}_D/|E|$ definierte Größe $\mu = 41\ cm^2V^{-1}s^{-1}$.

8.4.3.3. Deutung des Wiedemann-Franzschen Gesetzes. Besteht in einem Leiter der Wärmeleitfähigkeit λ ein Temperaturgradient $(\partial T/\partial x)$, so wird eine Wärmeenergie

(14) $E_w = \lambda(\partial T/\partial x)\, dA\, dt$

in der Zeiteinheit dt durch die Fläche dA in x-Richtung transportiert. Wird nun dieser Energietransport ausschließlich durch Leitungselektronen bewerkstelligt, so muß gelten: $E_w = n'E_{e,x} - n'E_{e,-x}$. Dabei sollen n'-Teilchen der Energie $E_e(x)$ in positiver (x)- und negativer $(-x)$-Richtung fliegen. Ein bei dA ankommendes Teilchen hat seit seinem letzten Stoß im Mittel die Strecke l zurückgelegt. Die Energie E_e ist $E_e = (3/2)\, k_BT(x)$. Also ergibt sich die transportierte Wärmeenergie E_w als Differenz zweier Energieströme mit Teilchenenergien. $E_{e,x} = (3/2)\, k_BT(x - v\tau)$, $E_{e,-x} = (3/2)\, k_BT \times (x + v\tau)$ und es gilt durch Reihenentwicklung: $T(x \mp v\tau) \approx T(x) \mp \frac{\partial T}{\partial x} v\tau$. Die Zahl der durch dA hindurchfliegenden Teilchen n' ergibt sich aus der Ladungsträgerkonzentration n zu: $n' = (1/6)\, nv\, dA\, dt$. Schließlich wird also $E_w = (1/6)\, nv\, (3/2)\, k_B\, (\partial T/\partial x)\, 2v\tau\, dA\, dt$ und mit 8.4. (8)folgt: $E_w = (3nk_B^2T\tau/2m)\,(\partial T/\partial x)\, dA\, dt$. Der Vergleich mit 8.4.(14) ergibt

(15) $\lambda = (3nk_B^2T\tau/2m)$.

Das Wiedemann-Franzsche Gesetz findet man unter Einbeziehung von 8.4.(13) $\lambda/\sigma T = 3k_B^2/e^2 = 2{,}23 \cdot 10^{-8}\ V^2K^{-2} = L$.

Das ist eine gute Übereinstimmung mit dem experimentellen Wert. Dagegen ergaben sich Schwierigkeiten, mit demselben Herangehen die spezifische Wärmekapazität c der Metalle zu deuten. Nach dem Gleichverteilungssatz sollten die Leitungselektronen je Mol einen Anteil $c_e = \frac{\partial E}{\partial T} = \frac{\partial}{\partial T} \frac{3}{2} kTL$ liefern. Das würde $c_e = \frac{3}{2} R$, mit R der Gaskonstanten, bedeuten. Nach der Dulong-Petitschen Regel ergäbe sich $c_e \approx R$. Dieser Wert ist viel zu hoch. Eine Lösung dieses Widerspruchs erzielte SOMMERFELD durch Anwendung der Fermi-Dirac-Statistik auf das Elektronengas (↑ 8.2.4.).

8.4.3.4. Anwendung der Fermi-Dirac-Statistik auf Metalle. Die für Elektronen als Teilchen mit halbzahligem Spin geltende *Fermi-Dirac-Verteilungsfunktion* (↑ 8.2.2.2.) lautet:

(16) $f(E) = \{\exp[(E - E_F)/k_BT] + 1\}^{-1}$.

Sie unterscheidet sich von der Verteilungsfunktion der Boltzmannstatistik nur um den Summanden 1 im Nenner. Da es sich hier im historischen Ablauf der Wissenschaftsentwicklung um ein zunächst ungewohntes Ergebnis handelte – die Boltzmannstatistik konnte z. B. in der Thermodynamik alles sehr gut beschreiben – wird ein System, das der Fermi-Dirac-Statistik unterliegt, als entartet bezeichnet. Ein solches *entartetes System* sind die Leitungselektronen in einem Metall oder auch in einem Halbleiter, der eine hohe Zahl von Verunreinigungen oder anderen Defekten enthält. Bei Leitungselektronen macht sich die Fermi-Dirac-Statistik auch makroskopisch am deutlichsten bemerkbar. Bei 0 K werden die etwa $2N$ Elektronenzustände, N ist die Zahl der Metallatome eines bestimmten Leiters, mit etwa N Leitungselektronen nach dem Pauliprinzip besetzt. Daraus folgt, daß das Leitungselektronengas eine große Nullpunktsenergie besitzt. Ein Teil der Elektronen hat am absoluten Nullpunkt eine relativ hohe Energie. Bei Temperaturen $T > 0$ K wird dem Elektronengas durch thermische Stöße Energie in der Größenordnung k_BT zugeführt. Das ist aber aufgrund des Pauliprinzips nur für Elektronen möglich, die dadurch auf ein noch freies energetisches Niveau gelangen können. Das gilt überraschenderweise nur für die Elektronen, die nahezu schon die höchste Energie besitzen (Abb. 8.4.3-1).

Mathematisch einfach läßt sich dieser Sachverhalt beschreiben, wenn wir das *Modell freier Elektronen* benutzen (↑ 8.2.2.1.). In diesem Fall wird das durch die positiven Atomrümpfe erzeugte periodische Potential als im ganzen

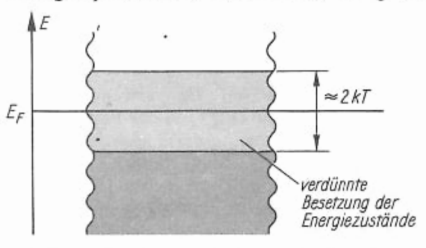


Abb. 8.4.3-1 Schematische Darstellung der thermischen Anregung von Leitungselektronen

Metallvolumen konstant, z. B. auch als $E_p = 0$, angesetzt. Dann kann man jedem Elektron die Energie

(17) $E = \boldsymbol{p}^2/2m_0 = \hbar^2\boldsymbol{k}^2/2m_0$

zuordnen.

Hier bedeuten $\boldsymbol{p}$ den Impuls und $\boldsymbol{k}$ den Wellenzahlvektor des Elektrons. Bekanntlich beansprucht jeder Elektronenzustand im Impulsraum p_x, p_y, p_z ein Volumen von h^3/V

(↑ 8.1.2.4.). V ist das Volumen, in dem sich die Elektronen aufhalten. Die *Zustandsdichte* $D(E)$ (Abb. 8.4.3-2) erhält deshalb die Form:

$$(18) \quad D(p)\,\mathrm{d}p = g\,\frac{V}{h^3}\,4\pi\hbar^2 k^2\,\mathrm{d}(\hbar k).$$

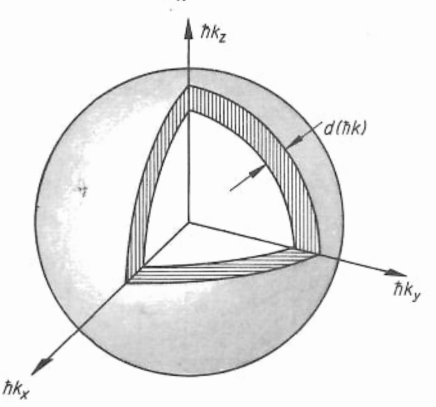


Abb. 8.4.3-2 Zur Berechnung der Zustandsdichte im Impulsraum

Der Faktor $g = 2$ berücksichtigt dabei die beiden möglichen Spinorientierungen auf einem Energieniveau.

$D(E)$ gibt die Zahl der Zustände der Elektronen an, die im Bereich E bis $E + \mathrm{d}E$ vorhanden sind. Aus 8.4.(17) und 8.4.(18) ergibt sich

$$(19) \quad D(E)\,\mathrm{d}E = V/2\pi^2(2m_0/\hbar^2)^{3/2}\,E^{1/2}\,\mathrm{d}E.$$

Hat man nun mit einer Elektronenkonzentration n im Volumen V zu rechnen, dann ergibt sich die maximale Elektronenenergie E_F aus:

$$nV = \int_0^{E_F} D(E)\,\mathrm{d}E;$$

$$nV = V/2\pi^2\,(2m_0/\hbar^2)^{3/2} \int_0^{E_F} E^{1/2}\,\mathrm{d}E;$$

$$n = \frac{1}{2\pi^2}\left(\frac{2m_0}{\hbar^2}\right)^{3/2}\frac{2}{3}\,E_F^{3/2} = \frac{2}{3}\,\frac{D(E_F)\,E_F}{V}.$$

$$(20) \quad E_F = (\hbar^2/2m_0)\,(3\pi^2 n)^{2/3}.$$

Die am absoluten Nullpunkt höchste Energie eines Leitungselektrons ist die Fermienergie E_F. Geht man hier wieder wie im Abschnitt 8.4.3.2. von $n = 8{,}5 \cdot 10^{22}\,\mathrm{cm}^{-3}$ der Elektronenkonzentration des Kupfers aus, so ergibt sich die für viele Metalle charakteristische Größenordnung der *Fermienergie* $E_F = 6{,}9$ eV. Setzt man diesen Wert in Formel 8.4.(16) ein und prüft, ab wann sich die 1 im Nenner vernachlässigen läßt, dann ergibt sich für Metalle eine praktisch unmöglich hohe Temperatur: $T_F = E_F/k_B = 80 \cdot 10^3$ K. Erst oberhalb $80 \cdot 10^3$ K wäre es möglich, Leitungselektronen nach der Boltzmannstatistik zu behandeln, weil dann auch alle im Leitungsband tiefliegenden Energiezustände durch einen einzigen Stoß entleert werden könnten.

Die Konzentration der Elektronen im Leitungsband bei $T > 0$ K ist nach 8.4.(20) durch

$$(21) \quad n = 1/V_0 \int_0^{\infty} D(E) f(E)\,\mathrm{d}E$$

gegeben. Die Integrationsgrenzen reichen eigentlich von der Unterkante des Leitungsbands $E_c = 0$ bis zu dessen Oberkante $E_c + \Delta E_c$. Man kann aber die obere Grenze gleich unendlich setzen, weil $f(E)$ bei $E > E_F$ sehr schnell gegen Null geht. Wenn man noch einmal die in guter Näherung geltende Beziehung 8.4.(20) heranzieht, läßt sich die mittlere Energie eines Leitungselektrons, bezogen auf die Unterkante des Leitungsbands berechnen.

Man findet mit den auch bisher verwendeten Zahlenwerten $\bar{E} = 4{,}1$ eV und $T_e = 48 \cdot 10^3$ K. Hierin liegen wesentliche Ursachen, daß mit der Anwendung der klassischen Elektronentheorie zwangsläufig Fehler entstehen.

8.4.3.5. Fermi-Dirac-Statistik und Elektronenanteil der spezifischen Wärmekapazität von Metallen. Da die Elektronenniveaus des Leitungsbands bis zur Fermienergie nahezu lückenlos besetzt sind, ist es leicht einzusehen, daß nur Elektronenniveaus in der Umgebung von E_F wirklich an der thermisch bewirkten Umsetzung teilnehmen können. Dieser Anteil beträgt

$$(22) \quad a \approx \int_{E_F}^{\infty} D(E) f(E)\,\mathrm{d}E/nV.$$

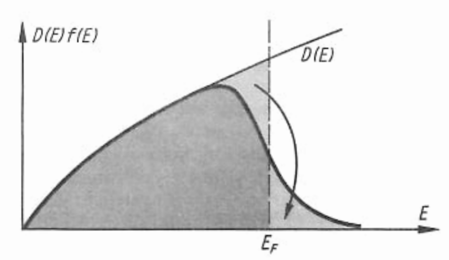


Abb. 8.4.3-3 Zur Berechnung des Elektronenanteils der spezifischen Wärmekapazität

Die Energie, die dieser Elektronenanteil aufgenommen hat, (Abb. 8.4.3-3) beträgt:

$$(23) \quad E_{\mathrm{th,e}} = \int_{E_F}^{\infty} 2(E - E_F)\,D(E) f(E)\,\mathrm{d}E.$$

Man kann diesen Energiewert grob annähern, wenn man berücksichtigt, daß die Zustands-

dichte in der Umgebung von E_F zu nehmen ist.

$$E_{th,e} = 2D(E_F) \int_{E_F}^{\infty} (E - E_F) \frac{dE}{\exp\left(\frac{E - E_F}{k_B T}\right) + 1}$$

und mit $x = (E - E_F)/k_B T$

wird daraus $E_{th,e} = 2D(E_F) \int_0^{\infty} \frac{x(k_B T)^2}{\exp x + 1} dx.$

Die Lösung dieses bestimmten Integrals kann einer Integraltabelle entnommen werden und man findet

(24) $E_{th,e} = 2D(E_F)(k_B T)^2 (\pi^2/12).$

Daraus folgt mit Hilfe von 8.4.(20) $E_{th,e} = \frac{\pi^2}{4} \frac{(k_B T)^2}{E_F} nV$ und die spezifische Wärmekapazität der Leitungselektronen für ein Mol, d. h. $nV = L$:

(25) $C_e = \frac{\partial E_{th,e}}{\partial T} = \frac{\pi^2}{3} \frac{k_B T}{E_F} \frac{3}{2} R.$

Vergleicht man dieses Ergebnis mit dem nach der klassischen Elektronentheorie erhaltenen, so zeigt sich, daß nur der Bruchteil $(\pi^2 k_B T)/3E_F$ der Elektronen an der Übernahme von Wärmeenergie bei Erwärmung beteiligt ist. Für Zimmertemperatur findet man, daß 1,2% der Leitungselektronen thermisch angeregt werden oder man erhält in Übereinstimmung mit dem Experiment etwa 1% des klassisch berechneten Werts.

Obwohl dieser Anteil an der spezifischen Wärmekapazität der Metalle sehr klein ist, kann er bei tiefen Temperaturen gut ermittelt werden. In guter Näherung gilt $C = C_G + C_e$ und der Gitteranteil geht mit $C_G \sim T^3$ gegen Null, so daß von einer bestimmten Temperatur abwärts $C_e > C_G$ wird, weil $C_e \sim T$ viel schwächer abnimmt.

8.4.3.6. Elektronenaustritt aus Metallen. Es ist eine völlig geläufige Tatsache, daß unter normalen Bedingungen die Leitungselektronen ein Metall nicht verlassen können. Allerdings gibt es eine ganze Reihe von experimentellen Möglichkeiten, wie Elektronen aus einer Metalloberfläche befreit werden können. Das sind oft gleichzeitig sehr effektive Methoden zur Konstruktion von Elektronenquellen.

Wird ein Metall sehr stark erhitzt, dann kommt es zur thermischen Elektronenemission. Die erreichbare Stromdichte j ergibt sich aus der *Richardson-Dushman-Gleichung*

(26) $j = AT^2 \exp(-E_A/k_B T).$

E_A ist eine Austrittsarbeit, die mindestens aufgebracht werden muß, damit ein Elektron aus einer Metalloberfläche austreten kann. Die Austrittsarbeit für Metalle besitzt Werte von 1,9 bis 5,3 eV (Tab. 8.4.3-4).

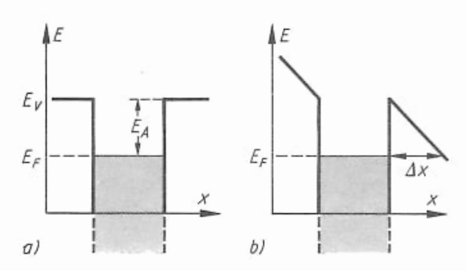


Abb. 8.4.3-5 Potentialtopfmodell der Metalle; a) Elektrisches Feld = 0, b) Elektrisches Feld ≠ 0; E_V Vakuumpotential, E_A Austrittsarbeit

Sie ist sehr stark von der Oberflächenbeschaffenheit abhängig. Ihre sehr starke Erniedrigung durch Fremdschichten wird technisch für Glühkatoden ausgenutzt. Man erreicht z. B. durch eine Bariumoxidschicht auf Platin 1,0 eV oder eine Bariumschicht auf Wolfram 0,30 eV. Da A für die meisten Metalle im Bereich 120 bis 200 $Am^{-2}K^{-2}$ liegt, wird laut Gleichung 8.4.(26) durch Absenkung der Austrittsarbeit eine starke Zunahme der Emissionsstromdichte erreicht. Glühkatoden hoher Lebensdauer sind vor allem in Elektronenröhren erforderlich. Elektronenquellen, meist auf der Basis von thoriumhaltigem Wolfram gefertigt, sind überall dort erforderlich, wo im Vakuum oder in Gasen mit freien Elektronen gearbeitet werden soll. Das ist z. B. der Fall in Fernsehröhren, in Fernsehaufnahmeröhren, in Elektronenmikroskopen, in Elektronenstrahlbearbeitungseinrichtungen und bei Gasentladungsleuchten. Will man mit derartigen thermisch emittierten Elektronen, z. B. in Form eines Elektronenstrahls arbeiten, dann müssen diese sehr niederenergetischen

Tab. 8.4.3-4 Austrittsarbeit W_A von Metallen in eV

W_A	5,3	4,7	4,6	4,5	4,4	4,3	4,2	4,1	4,0	3,7	3,2	2,5	2,3	2,2	1,9
Metall	Pt	Fe	Ni Cr Rh	W Cu	Au Sn	Co Zn	Mo Al Ta	Ag	Pb Cd	Mg	Ca	Ba	Na	K	Cs

Elektronen in einem elektrischen Feld nachbeschleunigt werden. Das ist nur in einem guten Vakuum möglich, da die mittlere freie Weglänge niederenergetischer Elektronen in Gasen sehr klein ist. Die Richardson-Dushman-Gleichung macht die wichtige Aussage, daß auch bei niedrigen Temperaturen eine sehr geringe Zahl von Elektronen ständig emittiert wird. Setzt man z. B. $A = 200\ \mathrm{Am^{-2}K^{-2}}$, $T = 300$ K und $E_A = 4$ eV, dann ergibt sich die Emissionsstromdichte $j = 5{,}86 \cdot 10^{-67}\ \mathrm{Acm^{-2}}$. Das heißt pro $\mathrm{cm^2}$ Metalloberfläche wird in 10^{40} Jahren nur ein Elektron emittiert. Ein Leiter verliert also durch Glühemission bei Zimmertemperatur überhaupt keine Elektronen. Führt man die Rechnung allerdings mit $E_A = 0{,}5$ eV aus, ergeben sich $j = 3{,}71 \cdot 10^{-6}\ \mathrm{Acm^{-2}}$ oder $2{,}32 \cdot 10^{13}$ Elektronen pro Sekunde. Dieser Spezialfall beschreibt pauschal recht gut den Kontakt zweier Metalle, zwischen denen sehr wohl kurzzeitig ein kräftiger Elektronenübertritt stattfinden kann.

Ein recht gutes Bild kann man sich von den Potentialverhältnissen in einem räumlich begrenzten Leiter machen, wenn man Abb. 8.4.3-5a zu Rate zieht. Wir haben hier im Inneren des Metalls ein konstantes Potential $E_P = 0$ angesetzt. Die Vernachlässigung der Gitterperiodizität macht sich hier nicht stark bemerkbar, da wir sowieso davon ausgehen, daß E_F räumlich konstant ist. Die Nachlieferung von Elektronen für die Glühemission erfolgt immer aus dem Elektronenvorrat des Elektronengases, der Energien $E \approx E_F \pm k_B T$ besitzt.

Bekanntlich entstehen an Metallen örtlich sehr hohe elektrische Feldstärken, wenn der Krümmungsradius der konvexen Oberfläche sehr klein ist. In Luft erfolgt bei hohen Spannungen ein elektrischer Durchschlag, verbunden mit einer Funkenentladung. Der zugrunde liegende Mechanismus wird als Feldemission bezeichnet. An der Oberfläche eines Golddrahts von 20 µm Durchmesser, der gegenüber seiner im Mittel 1 mm entfernten Umgebung eine Spannung von 10 V besitzt, herrscht etwa eine elektrische Feldstärke von $|E| = 10^3\ \mathrm{Vcm^{-1}}$. Der Potentialverlauf hat dann die in Abb. 8.4.3-5b skizzierte Form. Die Breite der Potentialbarriere der Höhe $\Delta E = E_A$ für Leitungselektronen mit maximaler Energie E_F ist $\Delta x \approx E_A\, eE$.

Für den hier betrachteten Fall ist $\Delta x = 44$ µm. In einem solchen in der Elektronik häufig auftretenden Beispiel kommt es nicht zu Feldemission. Gelingt es aber durch Anlegen weit höherer Spannungen oder durch Veränderung der Geometrie, z. B. durch Verwendung einer Spitze die elektrische Feldstärke an der Metalloberfläche stark anzuheben, kommt es zu Feldemission. Liegt zum Beispiel eine Feldstärke von $10^6\ \mathrm{Vcm^{-1}}$ vor, dann wird $\Delta x = 44$ nm und der Elektronenaustritt durch Tunneleffekt gewinnt an Bedeutung. Die Feldemissionsstromdichte wird unter Verwendung eines dreieckförmigen Potentialwalls in der WKB-Tunnelnäherung (↑ 3.3.); B Konstante:

(27) $j = B/E_A \exp(-4E_A/3\hbar\, eE)$.

Weitere viel verwendete Methoden des Elektronenaustritts sind der äußere Fotoeffekt und die Sekundärelektronenemission. Um eine meßbare Elektronenemission von einer Metalloberfläche zu erreichen, müssen Photonen mit einer Energie $h\nu > E_A$ auf das Metall auftreffen. In Fotozellen wird dieser Effekt zum Nachweis von Lichtintensitäten benutzt. In Sekundärelektronenvervielfachern wird diese spezielle Elektronenemission dadurch weiter verstärkt, daß man die aus der Fotokatode austretenden Elektronen im Vakuum in einem elektrischen Feld beschleunigt und etwa 20mal auf eine Elektrode aufprallen läßt. Die Verstärkung erfolgt durch Sekundärelektronenemission. Aus Metalloberflächen können Elektronen sowohl durch aufprallende Elektronen als auch durch Ionen und allgemein durch beschleunigte Teilchen befreit werden. Heute wird dieser Effekt auch vielfach zur Festkörperdiagnostik verwendet, indem man die räumliche Verteilung der z. B. durch einen gerasterten Teilchenstrahl erzeugten Sekundärelektronen auswertet.

8.4.4. Bändermodell bei Metallen

8.4.4.1. Brillouinzonen in Metallen. Der periodische Gitteraufbau von Metallen führt zwangsläufig zu einem periodischen Gitterpotential. Geht man mit einem angemessenen Potentialansatz, z. B. dem periodischen δ-Potential des Kronig-Penney-Modells an die Lösung der Schrödingergleichung heran, so erhält man unter Beachtung der periodischen Randbedingungen wie bei den Halbleitern eine Bänderstruktur der erlaubten Energieniveaus. Im Gegensatz zu den Ergebnissen der Theorie freier Elektronen erstrecken sich die Zustandsdichteverläufe $D(E)$ nicht lückenlos über einen großen Energiebereich, sondern bestimmte Energiegebiete können prinzipiell nicht stationär durch Elektronen eingenommen werden.

Die Bandstruktur $E(\boldsymbol{k})$ läßt sich heute für Metalle auf ± 1 meV genau berechnen. Eine gute Näherung liefert unter anderem die OPW-Methode. Sie verwendet orthogonalisierte ebene Wellen einer mathematisch schon recht komplizierten Form. Die OPW-Wellenfunktionen werden durch Kombination der ebenen Wellen quasifreier Leitungselektronen mit den Wellenfunktionen der lokalisierten Elektronen des Atomrumpfes gewonnen. Jedes Band kann

sukzessive mit Elektronen besetzt werden. Eine grobe Übersicht, ob ein metallischer Zustand vorliegt, kann auch rein theoretisch dadurch gefunden werden, daß man untersucht, wieviele Leitungselektronen für das energetisch am höchsten liegende Band zur Verfügung stehen. Ist die in einem Metallkristall vorhandene Zahl der Elementarzellen N, die Zahl der Basisatome q und z die Zahl der Elektronen in der Außenschale des Metallatoms, dann folgt als Bedingung für ein nicht vollständig gefülltes Band: $qzN = (2n - 1) N$, $n = 1,2$

(28) $qz = 2n - 1$.

Da sich die Elektronen in den Valenzbändern und im Leitungsband im energetisch tiefstmöglichen Zustand einbauen, ist bei der Überlappung von zumeist zwei Bändern Bedingung 8.4.(28) noch leichter erfüllbar. In Abb. 8.4.4-1 sind insbesondere die Verhältnisse hervorgehoben, die bei Überlappung eines an sich bis zum Rand gefüllten Bands mit einem vorher leeren Band auftreten.

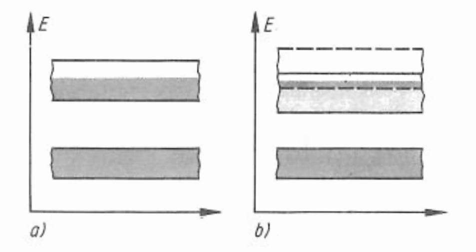


Abb. 8.4.4-1 Bändermodell und Zustandsdichte bei Metallen; a) isoliertes Leitungsband, b) Bandüberlappung eines völlig besetzten Bands mit einem leeren Band

Die hohe metallische Leitfähigkeit ist darauf zurückzuführen, daß sich die Elektronen in teilweise gefüllten Bändern unter dem Einfluß eines äußeren Felds gerichtet bewegen können. Der wesentliche Unterschied zu den Halbleitern besteht darin, daß dieser Leitungselektronenvorrat auch schon bei 0 K bereitsteht. Die grundlegenden Gleichungen des Modells freier Elektronen können formal beibehalten werden, wenn man auch bei den Metallen die Elektronenmasse durch die *effektive Masse* m^* (↑ 8.2.3.3.) ersetzt. Allerdings ist hier das Leitungsband so stark besetzt, daß keinesfalls die Darstellung vom m^* als eine für alle Leitungselektronen gleiche Größe angesehen werden kann.
In nahezu allen technisch wichtigen Fällen werden Leiter in polykristalliner, makroskopisch isotroper Form angewandt. Deshalb kann m^* näherungsweise als ein skalarer Mittelwert über alle Richtungen des Winkels zwischen elektrischem Feld und Kristallorientierung einerseits und über alle Energieniveaus dargestellt werden. Das entspricht einer Mitteilung über den dreidimensionalen $\boldsymbol{k}$-Raum.
Der Gesamtvorrat eines Leitungsbands an Zuständen für Elektronen wird durch die *Zustandsdichte* beschrieben. Die stationär erlaubten Energiezustände füllen im $\boldsymbol{k}$-Raum ein Volumen aus, dessen Begrenzung dem am höchsten liegenden Energieniveau $E_{max}(\boldsymbol{k})$ entspricht. Dabei schließt die Fläche $E_{max}(\boldsymbol{k})$ die *1. Brillouinzone* ab (↑ 8.1.2.4.). Durch neueste Forschungsergebnisse (↑ 8.4.1.4.) sind gerade die bisher vorwiegend für die Abzählung der Zustände benutzten Brillouinzonen für eindimensionale und zweidimensionale Leiter physikalisch bedeutsam geworden.
Nach der Vorschrift: Bestimmung des Gitteraufbaus, Konstruktion des reziproken Gitters (↑ 8.1.2.4.) und Ermittlung des kleinsten Volumens um den Ursprung, das sich finden läßt, wenn man senkrecht zu $\boldsymbol{G}/2$ stehende Ebenen zum Schnitt bringt, wird die 1. Brillouinzone gefunden. Die G-Vektoren sind Verbindungslinien zwischen beliebigen Punkten im reziproken Gitter. *Höhere Brillouinzonen* erhält man, indem man den letzten Verfahrensschritt wiederholt, aber sukzessive das sich aus diesem Konstruktionsprinzip ergebende nächstgrößere Volumen aufsucht. Man versteht das am besten anhand der Abb. 8.4.4-2 und 8.4.4.-3. In Abb. 8.4.4-4 sind einige dreidimensionale Brilluoinzonen zusammengestellt. Die Zahl der Zustände in einer einzelnen Brillouinzone beträgt $2Nq$. Nq ist die Anzahl der in einem Leiter vorhandenen Atome.

8.4.4.2. Fermiflächen in Metallen. Welche der $2Nq$-Elektronenzustände eines Leitungsbands wirklich besetzt werden, regelt die *Fermi-Dirac-Statistik*. Die Besetzung erfolgt bei $T = 0$ K so, daß alle Zustände bis $E(\boldsymbol{k}) = E_F$ aufgefüllt werden. Im $\boldsymbol{k}$-Raum ist $E(\boldsymbol{k}) = E_F$ eine Fläche, die Oberfläche des *Fermikörpers*. Sie wird kurz als Fermifläche bezeichnet. Abb. 8.4.4-5 charakterisiert den Zusammenhang zwischen Fermifläche und Brillouinzone näher. Im zweidimensionalen Beispiel schrumpft die Begrenzung des *Fermikörpers*

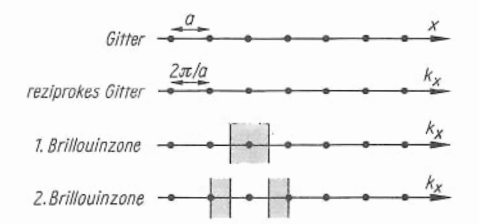


Abb. 8.4.4-2 Zur Konstruktion der Brillouinzonen in einem eindimensionalen Leiter

Abb. 8.4.4-3 Zur Konstruktion der Brillouinzonen im zweidimensionalen Leiter; a) Gitter, b) reziprokes Gitter, c) 1. Brillouinzone, d) 2. Brillouinzone

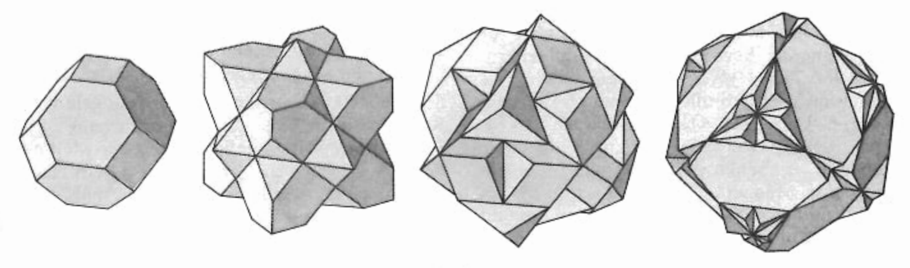


Abb. 8.4.4-4 Erste vier Brillouinzonen des kfz-Gitters

auf eine Linie zusammen. Im ersten Fall $a = b$ paßt die *Fermifläche* ganz in die Brilluoinzone hinein. Im zweiten Fall reicht der Fermikörper bis an die Zonengrenze heran. Im wiederholbaren Zonenschema – jeder Gitterpunkt des reziproken Gitters ist wahlweise Ursprung – entstehen in k_x-Richtung offene Strukturen. Derartige, häufig auftretende Fermikörper werden von *offenen Fermiflächen* begrenzt. Ein dreidimensionales Beispiel dafür ist die Fermifläche des Kupfers (Abb. 8.4.4-6).

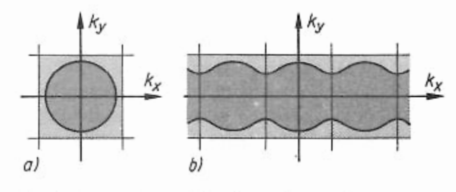


Abb. 8.4.4-5 Fermifläche in einem zweidimensionalen Metall; es wurde ein Valenzelektron je Atom angenommen; a) $a = b$, b) $2a = b$

Bei Temperaturerhöhung $T > 0$ K ändert sich die Form einer Fermifläche grundsätzlich nicht. Nur die Elektronen, die sich auf einer Haut der Dicke $\approx \pm 2k_BT$ um die Fermifläche herum aufhalten, werden durch thermische Anregung, elektrische und magnetische Felder primär beeinflußt.

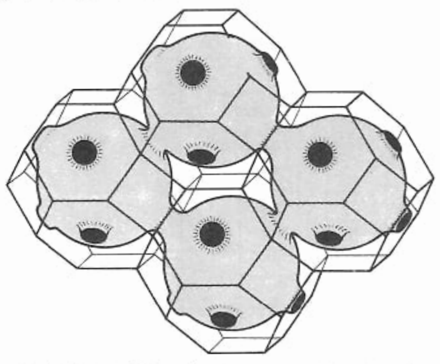


Abb. 8.4.4-6 Fermifläche des Kupfers im wiederholbaren Zonenschema

Nur sie finden freie Niveaus vor, die ihnen eine Energieaufnahme gestatten. Bei Zimmertemperatur ist also die Fermifläche eine rauhe Oberfläche, vergleichbar z. B. mit einer Kugel von 10 cm Radius mit einer Rauhtiefe von 80 μm. Es ist deshalb möglich, durch Messungen bei Zimmertemperatur die Fermifläche zu rekonstruieren. Das muß an hochreinen Metalleinkristallen geschehen. Die hier zu diskutierenden Methoden sind natürlich auch geeignet zur Bestimmung der Fermiflächen von Halbleitern. Mit Hilfe der *Zyklotronresonanz* kann die Krümmung der Fermifläche ermittelt werden. Man mißt dabei die Absorption einer Mikrowellenstrahlung von 2 bis 4 mm Wellenlänge durch den Probenkörper, der sich in einem homogenen Magnetfeld der Induktion B befindet. Stimmen die Mikrowellenfrequenz ω und die Zyklotronfrequenz $\omega_c = eB/m_c^*$ überein, so kommt es zur Resonanzabsorption.

Die *magnetische Widerstandsänderung* reagiert besonders stark auf offene Bahnen. Gemeint sind damit Elektronenbewegungen auf der Fermifläche, die mindestens zwei Fermikörper berühren. Eine besonders starke Änderung des Widerstands erhält man durch ein Magnetfeld, wenn die elektrische Feldstärke an der Probe mit der generellen Richtung einer offenen Bahn zusammenfällt. Kristallorientierungsabhängige Maxima der magnetischen Widerstandsänderung können also auf eine offene Fermifläche hinweisen und gestatten es, die sogenannten Hälse des Fermikörpers zu finden (↑ Abb. 8.4.4-6).

Beim *De-Haas-van-Alphen-Effekt* wird die magnetische Suszeptibilität gemessen. Man findet in einem hohen homogenen Magnetfeld eine periodische Folge von Maxima und Minima der Suszeptibilität. Aus der Differenz zweier aufeinanderfolgender Maxima bei B_n und B_{n+1} läßt sich mit $\Delta \frac{1}{B} = \left| \frac{1}{B_n} - \frac{1}{B_{n+1}} \right| = e/\hbar A_{\text{extr.}}$ die Extremalfläche $A_{\text{extr.}} \approx \pi k_{\text{extr.}}^2$ ermitteln. Extremale Querschnitte des Fermikörpers markieren sich deshalb so stark, weil ihre Zahl wesentlich höher liegt als die anderer Bahnen. Das läßt sich vielleicht am besten so verdeutlichen, daß man einen einfachen Vergleich heranzieht. Schneidet man einen Apfel von 5 cm Durchmesser in 1 mm dicke Scheiben und ordnet von oben beginnend diese nach Größen, läßt aber einen Spielraum von 10 mm zu, dann gehören die meisten Scheiben zur größten Klasse. In diesem Beispiel findet man in den fünf Größenklassen 4 bis 5 cm, 3 bis 4 cm usw. 60, 20, 12, 6 bzw. 2 Scheiben. Auf den Ausgangspunkt zurückgehend heißt das: durch die Elektronen auf und in der Nähe einer Extremalbahn wird ein hoher Prozentsatz der Elektronen auf der Fermifläche überhaupt repräsentiert.

Extremaldurchmesser der Fermifläche können auch aus dem periodischen Gang der Schallabsorption unter gleichzeitiger Einwirkung eines Magnetfelds gefunden werden. Das ist eine Anwendungsmöglichkeit des *magnetoakustischen Effekts*.

Ähnliches gilt für den *Gantmachereffekt*, bei dem die Änderung der Absorption von Radiowellen im MHz-Bereich unter der Wirkung eines parallel zur Metalloberfläche gerichteten homogenen Magnetfelds ausgewertet wird.

Schließlich kann auch durch eine sehr gelungene Kombination von Elementarteilchen- und Festkörperphysik, nämlich mittels der Methode der *Positronenannihilation*, die Fermifläche rekonstruiert werden. Bestrahlt man eine Metallprobe mit Positronen (↑ 3.1.), dann werden diese schnell durch Stöße abgebremst. Sie nehmen thermische Energien an, weil bei Energiewerten $E \approx k_BT$ für sie genügend freie Energieniveaus vorhanden sind. Da die positiven Atomrümpfe die Positronen abstoßen, vereinigen sie sich in einem Vorgang der Paarzerstrahlung bevorzugt mit den Leitungselektronen der Fermifläche. Der Impuls des Leitungselektrons und damit sein Wellenzahlvektor kann aus der genauen Analyse der zumeist paarweise entstehenden γ-Quanten entnommen werden. Dazu sind Messungen der Winkelkorrelationen gleichzeitig emittierter γ-Quanten erforderlich. Die Bestimmung einer Fermifläche wird immer in wechselseitigem Einsatz experimenteller Methoden und theoretischer Berechnungen zum Ziel führen.

8.4.4.3. Leitfähigkeit und Fermifläche. Es ist leicht einzusehen, daß durch ein angelegtes elektrisches Feld die Besetzung der Elektronenniveaus geändert wird. Das soll zunächst wieder für den reibungsfreien Fall des idealen Gitters betrachtet werden. Im feldfreien Fall hat der Fermikörper eine symmetrische Lage um den Ursprung des $\boldsymbol{k}$-Raums. Kristallographisch ausgedrückt könnte man sagen, der Ursprung des reziproken Gitters ist Inversionszentrum des Fermikörpers. Deshalb findet man zu jedem Elektron mit dem Quasiimpuls $\hbar\boldsymbol{k} = m^*\boldsymbol{v}$ ein Elektron mit dem Quasiimpuls $-\hbar\boldsymbol{k} = -m^*\boldsymbol{v}$ oder anders ausgedrückt, es gilt für den gesamten Fermikörper:

$$(29) \quad \sum^{nV} \hbar\boldsymbol{k} = 0$$

oder

$$\sum^{nV} m_i^* \boldsymbol{v}_i = 0\,.$$

Beim Anlegen eines elektrischen Felds an ein Metallstück mit der Leitungselektronenkonzentration n und dem Volumen V wirkt auf jedes Elektron die Kraft $-\mathrm{e}\boldsymbol{E}$. Jeder Elektronenimpuls ändert sich also pro Zeiteinheit um

$$(30) \quad \frac{\mathrm{d}(\hbar\boldsymbol{k})}{\mathrm{d}t} = -\mathrm{e}\boldsymbol{E}.$$

Verwendet man jetzt wieder die Überlegungen von Abschnitt 8.4.3.2., berücksichtigt aber die Bandstruktur durch $m = m^*(\boldsymbol{k})$, dann ergibt sich mit 8.4.(10) bis 8.4.(11):

$$(31) \quad \mathrm{d}(\hbar\boldsymbol{k}) = -\mathrm{e}\boldsymbol{E} = \mathrm{d}(m^*\boldsymbol{v}) = m^*\boldsymbol{v}_{\mathrm{Drift,\,max}}\,.$$

Hier wurde wieder davon ausgegangen, daß nur die mittlere freie Flugzeit für die Beschleunigungsphase zur Verfügung steht. Mit 8.4.(11) bis 8.4.(13) ergibt sich wieder der Zusammenhang mit den makroskopisch meßbaren Größen.

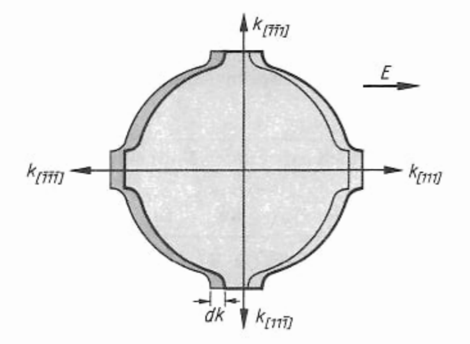


Abb. 8.4.4-7 Verschiebung des Fermikörpers unter Einwirkung eines elektrischen Felds; es ist die Verschiebung für einen Kupfereinkristall bei einer Feldrichtung parallel [111] dargestellt

Es sei aber ausdrücklich darauf hingewiesen, daß die effektive Masse eine Tensorgröße ist und damit auch σ oder ϱ Eigenschaften eines Tensors besitzen. In Abb. 8.4.4-7 ist der Schnitt durch einen Fermikörper dargestellt und seine Verschiebung stark übertrieben angedeutet. Normalerweise würde gelten $v_{\mathrm{Drift}}/v_{\mathrm{F}} \approx |\mathrm{d}\boldsymbol{k}|/|\boldsymbol{k}_{\mathrm{F}}| \approx 1{,}4 \cdot 10^{-9}$.
Dabei ist v_{F} nach $v_{\mathrm{F}} = (2E_{\mathrm{F}}/m_{\mathrm{F}}^*)^{1/2}$ und v_{Drift} mit maximal zulässiger elektrischer Feldstärke für Kupfer berechnet worden. Auf ein mechanisches Modell übertragen würde das bedeuten, daß ein Körper von 1 km Maximaldurchmesser als Abbild des Fermikörpers nur um rund 1 μm aus seiner Ruhelage verschoben wird.

8.4.4.4. Brillouinzonen und Fermiflächen konkreter Leiter. Bestimmte Eigenheiten des Elektronensystems eines Metallatoms, bzw. von mehreren Atomen bei Legierungen, übertragen sich auf die Elektronenniveaustruktur des Leiters. Alkalimetallatome haben ein s-Elektron (↑ 4.2.3.2.) als Valenzelektron. Die s-Zustände der Valenzelektronen verschmelzen zum halb besetzten Leitungsband. Die Fermifläche ist annähernd eine Kugeloberfläche. Außer bei Li ($m^* = 1{,}40\ m_0$) ist $m^*/m \approx 1$. Die Leitungselektronen der Alkalimetalle verhalten sich annähernd wie freie Elektronen. Dagegen haben die einwertigen Kupfer-, Silber- und Goldatome ein mit 10 Elektronen vollbesetztes $3d$-, $4d$- bzw. $5d$-Niveau und das nächsthöhere s-Niveau ist mit je einem Elektron besetzt. Aus diesen Elektronenniveaus baut sich im Festkörper ein einheitliches Leitungsband auf, zu dessen vollständiger Besetzung 12 Elektronen erforderlich wären. Der zugehörige Fermikörper weicht stärker von der Kugelgestalt ab, wie das Abb. 8.4.4-6 zeigt. Die effektiven Massen betragen auch hier recht genau $m^* \approx 1 m_0$. Zweiwertige Metalle wie Be, Zn und Cd müßten eigentlich ein vollbesetztes höchstes Band besitzen, denn ihr äußeres s-Niveau ist zweifach besetzt. Der metallische Zustand und nicht der eines Isolators ergibt sich hier aus der Überlappung des s-Bands mit dem nächsthöheren p-Band. Dadurch fehlen aber dann durch Umverteilung Elektronen im s-Band. Wie bei den Halbleitern muß man auch mit dem Auftreten von Defektelektronen rechnen. Diese lassen sich tatsächlich durch Halleffektmessungen nachweisen. Beim dreiwertigen Aluminium liegen wie bei anderen dreiwertigen Metallen genähert wieder Verhältnisse wie bei freien Elektronen vor.

Eine Sonderstellung nehmen die Übergangsmetalle ein. Sie besitzen durchweg unvollständig besetzte d-Bänder, die sich mit s-Bändern überlappen, während die f-Bänder leer oder voll besetzt sind. Dadurch entstehen nur teilweise besetzte Leitungsbänder.

Bei Legierungen sind logischerweise komplexere Zusammenhänge wirksam. Vieles kann bei weitem noch nicht so klar theoretisch beschrieben werden wie bei elementaren Metallen. An einem Beispiel soll aber gezeigt werden, wie das Auftreten eines bestimmten Gittertyps in einer Legierung in Abhängigkeit von der Konzentration auch von der Form der Brillouinzonen und des Festkörpers abhängen kann. Das ist sehr ausgeprägt bei den *Hume-Rothery-Phasen* (↑ 6.4.4.) der Fall. Bestimmte Legierungen einer Komponente A = Cu, Ag, Au, Ni, Pt usw. mit zweiwertigen bis fünfwertigen Elementen B = Zn, Hg, Sn, Si, Ca, Al usw. weisen beim Legieren von 0 bis 100 Atomprozent eine charakteristische Reihe von Modifikationen auf. Das sind die A-reiche α-*kfz*-Phase, die β-*krz*-Phase, die γ-Phase mit 52 Atomen in der kubischen Riesenzelle, die hexagonal dichteste ε-Phase und die B-reiche, hexagonale η-Phase. Gemeinsam ist allen auch noch so unterschiedlich zusammengesetzten Legierungen (z. B. γ: Cu_4Zn_8, Ni_5Zn_{21}, $Cu_{31}Sn_8$ usw.), daß das Hauptstabilitätskriterium einer bestimmten Modifikation die mittlere Valenzelektronenzahl n_e pro Legierungsatom ist. Man findet z. B. für die β-Phase $n_e = 3/2$ und für die γ-Phase $n_e = 21/13$. Eine dem wirklichen Sachverhalt sehr nahe kommende Erklärung geht davon aus, daß bei kritischen Valenzelektronenzahlen z. B. bei $n_e \approx 3/2$ der als kugelförmig angenommene Fermikörper die Brillouinzonengrenzen des betreffenden Gitters, z. B. des *kfz*-Gitters, gerade berührt, eine größere Fermikugel also nicht hineinpaßt. Erst der nächste Gittertyp, z. B. das *krz*-Gitter, ist dann mit seiner Brillouinzone wieder aufnahmefähiger. Zur Erklärung sei angemerkt, daß A immer das niederwertigere Element ist. Die räumliche Anordnung der Legierungsatome wird also auch durch die konkreten Zustandsdichteverläufe der Elektronen mitbestimmt.

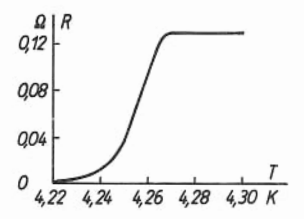


Abb. 8.5.1-1 Übergang von Quecksilber in den supraleitenden Zustand; R Widerstand, T Temperatur

8.5. Supraleiter

8.5.1. Supraleitende Stoffe

8.5.1.1. Kennzeichen eines Supraleiters. Das Phänomen der Supraleitung (SL) wurde 1911 von Kamerlingh-Onnes an Quecksilber entdeckt (Abb. 8.5.1-1). Während bei Normalleitung (NL) der spezifische Widerstand bei tiefen Temperaturen durch den Restwiderstand charakterisiert wird, sinkt bei Supraleitern der Widerstand auf einen unmeßbar kleinen Wert ab. Das geschieht bei der Sprungtemperatur T_k. Für die Abschätzung des spezifischen Widerstands eines Supraleiters eignet sich am besten die Langzeituntersuchung eines Induktionsstroms. Ein in einem Ring fließender Strom nimmt nach $I = I_0 \exp(-Rt/L)$ ab. R ist der Widerstand des Rings und L die Selbstinduktion. Der Strom I kann anhand seiner magnetischen Kraftwirkung gemessen werden. Derartige Versuche ergeben: $\varrho_{SL} \leqq 10^{-23}$ Ωcm.
Neueste Messungen sprechen allerdings dafür, daß es bei Abkühlung nicht generell nur ein einmaliges Umschalten NL → SL gibt. An Legierungen des Typs $ErRh_4B_4$ konnte erstmalig gezeigt werden, daß auch zwei Sprungpunkte NL → SL und SL → NL bei Abkühlung möglich sind. Magnetfelder reduzieren in definierter Weise die Sprungtemperatur, bis der

Tab. 8.5.1-2 Sprungtemperaturen T_K in K von metallischen Elementen unter Normaldruck nach Roberts

Metall	Nb	Tc	Pb	V	La	Ta	Hg	Sn	In	Sb
T_K	9,25	7,78	7,23	5,43	4,88	4,47	4,154	3,721	3,405	2,7
Metall	Tl	Re	Pa	Th	Al	Ga	Mo	Zn	Os	Zr
T_K	2,39	1,697	1,4	1,39	1,175	1,083	0,916	0,875	0,655	0,53
Metall	Cd	Ru	Ti	Ir	Be	W	Am			
T_K	0,52	0,493	0,39	0,14	0,026	0,0154	0,79			

supraleitende Zustand ganz verschwindet. Da jeder elektrische Strom ein Magnetfeld erzeugt, ist damit auch eine Begrenzung des durch einen Supraleiter fließenden Stroms gegeben. Bisher fordert keine Theorie den Übergang jedes Metalls in den supraleitenden Zustand. Allerdings ist die Suche nach sehr niedrigen Sprungtemperaturen schwierig, da sehr geringe, magnetisch wirkende Verunreinigungen schon ausreichen, die Entstehung des supraleitenden Zustands zu verhindern.

8.5.1.2. Supraleitende Elemente, Legierungen und Verbindungen. Von den metallischen Elementen ist eine stattliche Anzahl supraleitend. 1978 konnte Supraleitung für das erste Transuranelement Americium nachgewiesen werden. Bei radioaktiven Elementen treten zusätzliche Kühlprobleme auf, weil sich die Meßprobe durch ihre eigene Strahlung aufheizt (Tab. 8.5.1-2). Bisher konnte für $T \geqq 30$ mK keine Supraleitung bei Cu, Ag, Au, Pd, Pt, Rh, Li, Na, K, Rb, Mg, Ca und Sr nachgewiesen werden. Eine empirische Regel für das wahrscheinliche Auftreten von Supraleitung wurde von MATTHIAS 1955 angegeben, in die sich auch alle neueren Entdeckungen bis in die Gegenwart einordnen lassen (Abb. 8.5.1-3). Besonders hohe Sprungtemperaturen sind bei einer mittleren Valenzelektronenzahl pro Gitterbaustein von 4,5 zu erwarten. Ein zweiter günstiger Bereich befindet sich in der Umgebung von 6,5 Valenzelektronen pro Baustein. Das Diagramm ist so zu verstehen, daß jeweils nahezu die gesamte Peakfläche mehr oder weniger dicht von Meßpunkten real existierender Supraleiter besetzt ist. Das nach oben offene Maximum weist darauf hin, daß noch keine Theorie

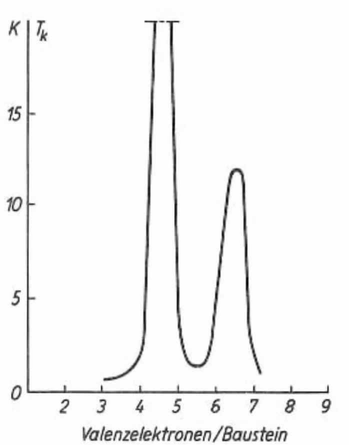


Abb. 8.5.1-3 Regel von MATTHIAS zur Auffindung von Supraleitern mit hoher Sprungtemperatur

die obere Grenze schlüssig zu berechnen gestattet. Nicht ausgeschlossen scheinen, vor allem mit Drei- und Mehrkomponentenverbindungen, 35 K zu sein. Bei vielen metallischen Elementen kann durch hohe Drücke (↑ 8.4.1.2.) die Sprungtemperatur verändert werden. Mit wenigen Ausnahmen, z. B. von Pb und Tl, steigt die Sprungtemperatur deutlich an. Allerdings werden in der Regel bei schon sehr hohen Sprungtemperaturen unter Normaldruck durch Druckerhöhung keine Verbesserungen erreicht.

Tab. 8.5.1-4 Sprungtemperatur T_K in K einiger ausgewählter Supraleiter

Typ	Supraleiter	T_K
Mehrkomponentensystem	GaV_3	14,6
	SiV_3	16,9
	Nb_3Sn	18,25
	Nb_3Al	18,8
	$Al_{0,85}Ge_{0,15}$	20,5
	$Nb_{2,85}Ta_{0,15}$	
Hochdruckmodifikation	P > 10 GPa	4,7
	Ge > 12 GPa	5,4
	Si > 12 GPa	7,1
	Pb II > 16 GPa	3,6
	Cs IV > 13 GPa	1,5
	La > 14 GPa	11,9
Dünnschichten technologisch optimiert	Nb	10,1
	Pb	7,7
	Mo	6,7
	Sn	6,0
	W	4,1
	Nb_3Ga, gesputtert	20
	Nb_3Ge, gesputtert	23
	$Nb_3Ge_xSi_y$, gesputtert	25
Schichtsysteme, Messung parallel zur Schichtebene	Al/TCNQ	5,24
	In/Anthrachinon	4,6
	NbS_2/Pyridin	4,0

Es ist auch nicht weiter überraschend, daß die metallischen Hochdruckmodifikationen z. B. des Phosphors, Germaniums und Siliciums supraleitend werden. Eher ist das beim Cs verwunderlich (Tab 8.5.1-4). Noch weitgehend ungeklärt sind aber die Gesetzmäßigkeiten, die zu einem Anstieg der Sprungtemperatur in Dünnschichten von Elementen und Mehrkomponentensystemen führen. Interessante Ergebnisse wurden auch erreicht durch die sandwichartige Schichtkombination von Metall und Isolator. Die überhaupt höchsten Sprungtemperaturen wurden an gesputterten Schichten des Nb_3Ge bzw. $Nb_3Ge_xSi_{1-x}$ mit etwa 25 K gemessen. Eine besonders geeignete Struktur

scheint die A 15- oder Cr_3Si-Struktur zu sein, in der solche intermetallischen Verbindungen wie Nb_3Sn, Nb_3Ge, Nb_3Al, GaV_3 und SiV_3 mit – am heutigen Stand gemessen – sehr hohen

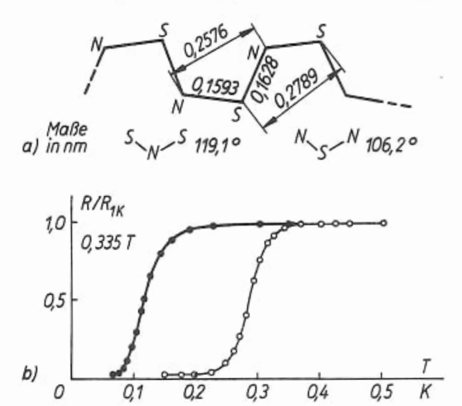


Abb. 8.5.1-5 a) Molekülstruktur des Polythiacyls $(SN)_x$ nach McDiarmid, b) NL-SL-Übergang von $(SN)_x$ nach Grenne, Street und Suter

Sprungpunkten kristallisieren. Läßt sich ein bestimmtes System wie Si/Nb nicht in dieser Form kristallisieren, treten auch keine besonders günstigen Sprungtemperaturen auf. 1975 wurde erstmalig auf völlig neuem Wege, nämlich mittels der kovalent gebundenen Polymerkette $(SN)_x$ Supraleitungen realisiert. Wenn auch deren Sprungtemperatur mit 0,26 bis 1,5 K sehr niedrig liegt, ist das doch ein echter Durchbruch zu einer neuen Supraleiterklasse (Abb. 8.5.1-5). 1980 wurden auch die ersten organischen Supraleiter, z. B. $(TMTSF)_2\,PF_6$ oder $(TMTSF)_2\,ClO_4$ gefunden. Ihre Sprungtemperaturen liegen bei etwa 1 K.

8.5.2. Supraleitender Zustand

8.5.2.1. Einteilung der Supraleiter. Man unterscheidet Supraleiter 1., 2. und 3. Art. Ein Supraleiter 1. Art oder die *Meißnerphase* des supraleitenden Zustands liegt vor, wenn ein bestimmtes Volumen eines Festkörpers homogen supraleitend ist. Zu den *Supraleitern 1. Art* gehören alle supraleitenden Elemente außer Nb, V und Tc. *Supraleiter 2. Art* können bei kleinem äußerem Magnetfeld zunächst eine Meißnerphase besitzen, müssen das aber nicht. Charakteristisch für den supraleitenden Zustand 2. Art, der auch als *Shubnikovphase* bezeichnet wird, ist das Vorhandensein von normalleitenden Gebieten in einer supraleitenden Matrix, also ein räumlich heterogener Zustand. Meist bilden sich die normalleitenden Zonen als sogenannte *Flußschläuche* aus (Abb. 8.5.2-1). Aufgrund von Anziehungskräften und Abstoßungskräften zwischen den Flußschläuchen ordnen sich diese im Gleichgewicht in einer zweidimensionalen dichtesten Packung. Wir beziehen hier den Terminus zweidimensional auf einen Querschnitt senkrecht zu den Achsen der Flußschläuche. Neuerdings wird auch der Begriff *Supraleiter 3. Art* verwendet. Hier handelt es sich um eine Shubnikovphase, in der durch Gitterdefekte die Flußschläuche fest verankert sind. Das ist wesentlich, wenn man verhindern will, daß bei einem Stromfluß durch den Supraleiter die Flußschläuche im Supraleitervolumen wandern. Eine Wanderbewegung der Flußschläuche infolge der Lorentzkraft bringt nämlich eine Erhöhung des spezifischen Widerstands mit sich.

8.5.2.2. Supraleiter im Magnetfeld. Wird ein Normalleiter in einem Magnetfeld abgekühlt, dann stellt man fest, daß der Übergang NL → SL bei einer tieferen Temperatur $T_K(H)$ erfolgt als ohne Magnetfeld, $H = 0$. Deshalb ist es immer möglich, den Übergang SL → NL auch isotherm durch Anlegen eines bestimmten äußeren Magnetfelds zu erreichen. Oberhalb einer *kritischen Feldstärke* $H_{k1}(T)$ geht die Meißnerphase entweder in den normalleitenden Zustand über (SL 1. Art) oder es entsteht die Shubnikovphase (SL 2. Art). Näherungsweise lassen sich die Meßkurven in Abb. 8.5.2-2 durch die Beziehung

$$(1) \quad H_{k1}(T) = H_{k1}(0)\,[1 - (T/T_K)^2]\,,$$

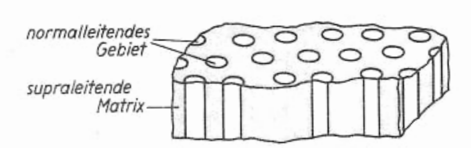


Abb. 8.5.2-1 Supraleiter zweiter Art mit Flußschläuchen

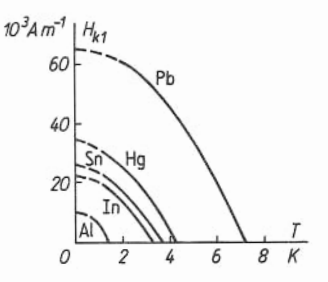


Abb. 8.5.2-2 Zustandsdiagramme verschiedener Metalle (Erläuterung im Text)

SL 1. Art NL ↔ SL bei $H_{k1}(T)$,
SL 2. Art SL 1. ↔ SL 2. bei $H_{k1}(T)$
darstellen. Der Übergang erfolgt bei Supraleitern 1. Art relativ scharf, nämlich innerhalb eines Temperaturintervalls von 50 mK. Der Übergang einer Shubnikovphase benötigt dagegen ein Temperaturintervall von 0,5 K.

Als *Meißner-Ochsenfeld-Effekt* bezeichnet man die Erscheinung, daß beim Übergang NL→ SL ein Magnetfeld aus dem Supraleiter herausgedrängt wird. Für das Innere eines Supraleiters 1. Art ist die magnetische Induktion $B = 0$ (Abb. 8.5.2-3).

Das kommt physikalisch dadurch zustande, daß in einer dünnen Oberflächenschicht des Supraleiters Ringströme fließen, die exakt den Feldeinfluß kompensieren. Damit ist aber ein Supraleiter 1. Art ideal diamagnetisch. Die magnetische Suszeptibilität ist $\varkappa_{mag} = -1$ oder die relative magnetische Permeabilität $\mu = 0$ (↑ 8.7.1.). Die Schichtdicke der Oberflächenschicht, in der ein äußeres Magnetfeld bis auf den e-ten Teil abgesunken ist, bezeichnet man als *Londonsche Eindringtiefe* oder kurz Eindringtiefe λ. Ist x der Abstand von der Oberfläche eines Supraleiters, dann gilt $x \gg \lambda$ für $B = 0$ und $E = 0$; $x \approx \lambda$ für $B = B_0 \exp(-x/\lambda)$.

Auch dieser charakteristische Parameter eines Supraleiters ist temperaturabhängig:

$$(2) \quad \lambda = \lambda_0 \left[1 - \left(\frac{T}{T_K}\right)^4\right]^{-1/2}$$

Diese Gleichung beschreibt in guter Näherung den Vorgang des durch ein Magnetfeld hervorgerufenen Übergangs SL → NL. Wird die Temperatur T bis auf T_K erhöht oder wird T_K durch ein geeignetes Magnetfeld bis T abgesenkt, ergibt sich $\lambda \to \infty$. Das Magnetfeld durchdringt dann den ganzen Festkörper, d. h., es liegt Normalleitung vor. Die Eindringtiefe λ_0 beträgt etwa 100 nm. Diese endliche Eindringtiefe führt zu *Anomalien bei Supraleitern*, die eine Abmessung dieser Größenordnung besitzen. Das ist der Fall bei dünnen Schichten unter $10\lambda_0$ Schichtdicke und bei der Shubnikovphase. Selbst wenn sich im letzteren Fall die Flußschlauchbereiche fast berühren, sind in den Zwischenräumen noch genügend supraleitende Fäden vorhanden. In beiden Fällen ergibt sich eine höhere kritische magnetische Feldstärke H_{k1} bzw. H_{k2}. Oberhalb H_{k2} erfolgt auch bei einem Supraleiter 2. Art der Übergang in die Normalleitung (Tab. 8.5.2-4). Wenn man ein Zustandsdiagramm im zweidimensionalen Zustandsraum (H, T) betrachtet, dann findet man für die Meißnerphase bzw. die Shubnikovphase die in Abb. 8.5.2-5 skizzierten Zustandsgebiete. Das Zustandsgebiet der Meißnerphase wird durch die H-Achse, die T-Achse und die Kurve $H_{k1}(T)$; die Shubnikovphase durch die H-Achse

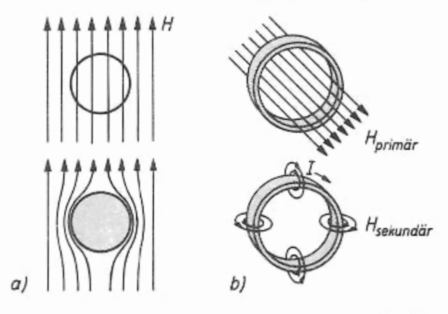


Abb. 8.5.2-3 Normalleiter und Supraleiter im Magnetfeld; a) Kugel, b) Ring

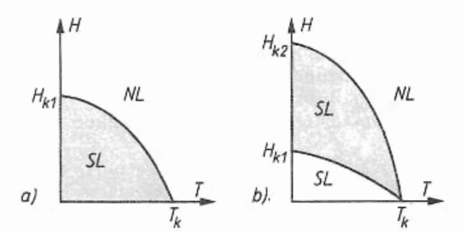


Abb. 8.5.2-5 Zustandsdiagramm; a) der Meißnerphase, b) der Shubnikovphase (Erläuterung im Text)

Tab. 8.5.2-4 Kritische magnetische Feldstärken H_{k1} und H_{k2} ausgewählter Supraleiter

Supraleiter 1. Art

Supraleiter	Al	Pb	Sn	In	W
H_{k1} in kAm^{-1}	8,30	0,41	2,44	2,26	0,091

Supraleiter 2. Art

Supraleiter	V	Nb	Nb_3Al	Nb_3Sn	$Nb_3Al_{0,75}Ge_{0,25}$
H_{k1} in kAm^{-1}	11,2	14,8	–	–	–
H_{k2} in kAm^{-1}	21,4	31,2	2,64	207	335

und die beiden Kurven $H_{k1}(T)$ sowie $H_{k2}(T)$ begrenzt. H_{k1} kann bei einem Supraleiter 2. Art auch Null sein (↑ Tab. 8.5.2-4), dann existiert überhaupt keine Meißnerphase.

8.5.2.3. Elektronenzustände in einem Supraleiter. Untersucht man die Sprungtemperatur von Leitern mit exakt gleicher Kristallstruktur und

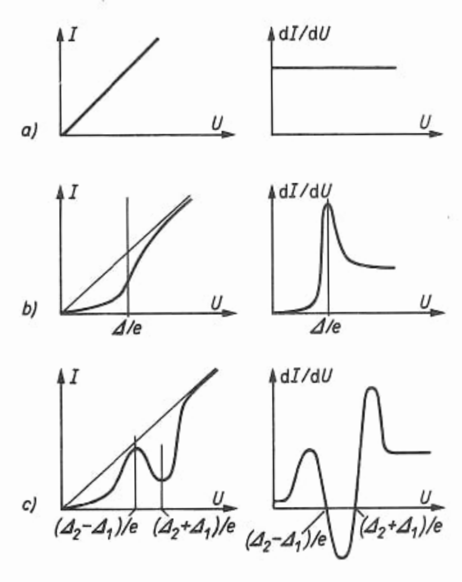


Abb. 8.5.2-6 Einelektronentunneleffekt in verschiedenen Grundstrukturen; a) Metall – Isolator – Metall, b) Supraleiter – Isolator – Metall, c) Supraleiter – Isolator – Supraleiter

mit gleicher Elementzusammensetzung, aber mit verschiedener Masse M der atomaren Bausteine, dann findet man den *Isotopieeffekt*. Das ist erfolgreich gemacht worden für Hg, Sn, Pb, Cd, Te, Zn ($n \approx 0,5$) und z. B. Mo, Os, Nb_3Sn ($n < 0,5$). Es ergibt sich die einfache Beziehung

$$(3) \quad T_K \sim M^{-n}.$$

Da nun auch die Debyegrenzfrequenz der Phononen und damit die Debyetemperatur der Relation $\Theta_D \sim M^{-1/2}$ folgt, ergibt sich $T_K \sim \Theta_D$. Das ist ein wertvoller Hinweis, daß die Supraleitung etwas mit einer Elektron-Phonon-Wechselwirkung zu tun hat.

Ein weiterer wichtiger Tatbestand kann gefunden werden, wenn man Elektronen durch eine 3 bis 10 nm dicke Isolatorschicht (I), z. B. Al_2O_3, tunneln läßt. Abgesehen von Feinheiten des Kennlinienverlaufs ist die Strom-Spannungs-Kennlinie linear, wenn der Tunnelübergang die Grundstruktur NL/I/NL hat. Ist aber die Sandwichprobe, z. B. in der Anordnung $Pb/Al_2O_3/Al$, soweit abgekühlt worden, daß ein Metall oder beide supraleitend geworden sind, dann findet man eine Nichtlinearität. Das ist in Abb. 8.5.2-6 dargestellt. Diese Experimente ergeben einen Hinweis auf die sogenannte *Energielücke eines Supraleiters* Δ. Mit Hilfe von Tunnelexperimenten kann Δ auch bestimmt werden. Obwohl ein Teil der Leitungselektronen Elektronenpaare bildet, bleibt in einem Supraleiter die Fermi-Dirac-Statistik sehr gut weiter anwendbar. Man kann deshalb prinzipiell aus Tunnelexperimenten die ungefähren Zustandsdichteverläufe der ungepaarten Elektronen ermitteln. Das ist allerdings auch auf anderem Wege z. B. durch Absorptionsmessungen mit Lichtquanten oder Ultraschall sowie theoretisch möglich und man erhält auch eine indirekte Aussage über die gepaarten Elektronen. Allerdings lassen sich die Elektronenpaare in Abb. 8.5.2-7 nicht mit unterbringen, weil sie der Bose-Einstein-Statistik genügen. Die *Tunnelkennlinien* der Abb. 8.5.2-6 gehen davon aus, daß ein geringer Teil der ungepaarten Elektronen bei $T > 0$ K thermisch angeregt ist. Es kann also auch bei sehr kleinen Spannungen ein Strom fließen. Ist nur die Energielücke eines einzelnen Supraleiters wirksam, dann kommt ein starker Stromanstieg erst dann zustande, wenn die durch das elektrische Feld bewirkte Anhebung der Elektronenniveaus um Δ erfolgt ist (Abb. 8.5.2-6b). Tunneln Elektronen aus dem Reservoir eines Supraleiters in das eines zweiten, dann erfolgt zunächst ein Stromanstieg, weil die thermisch angeregten Elektronen eine hohe Dichte unbesetzter Zustände auf der anderen Seite der Barriere vorfinden, wenn sie auf deren Niveau angehoben werden. Das erfordert eine Beschleunigungsspannung $U_1 = (\Delta_2 - \Delta_1)/e$. Der Tunnelstrom sinkt darauf bei steigender Spannung etwas ab, weil jetzt die Elektronen im Niveauspektrum des 2. Supraleiters eine kleinere Zustandsdichte vorfinden. Der Tunnelstrom steigt aber dann steil an, wenn $U_2 = (\Delta_2 + \Delta_1)/e$ erreicht ist. Hier steht dann für den Tunnelvorgang der entsprechend höhere Vorrat von Elektronen des Supraleiters 1 unterhalb des Ferminiveaus zur Verfügung (Abb. 8.5.2-6c). Als Erläuterung sei darauf hingewiesen, daß die im Normalleiter vorhandenen Zustände innerhalb der Energielücke sich beim Supraleiter unmittelbar in deren Umgebung, d. h. nahe der Fermienergie, zusätzlich einordnen. Anders ausgedrückt verlagern sich bei einem Übergang NL → SL, beginnend bei T_K, die Elektronenzustände der nächsten Umgebung der Fermifläche. Die Energielücke müßte man sich also im k-Raum als eine Schale der Dicke $dk \approx 2m\,\Delta/\hbar^2 k_F$ vorstellen. Unmittelbar innerhalb und außerhalb der Schale wird dagegen die Zustandsdichte beträchtlich erhöht. Es soll schon hier darauf hingewiesen werden, daß es

sich bei 2Δ um die Bindungsenergie der Elektronenpaare im Supraleiter (Meißnerphase) handelt. Aus der gründlichen Erforschung der Zustandsdichte der Normalkomponente können wertvolle Aussagen über die Elektronenpaare gewonnen werden. In Tab. 8.5.2-8 sind einige Energielückenwerte zusammengestellt.

8.5.2.4. Weitere wichtige Eigenschaften von Supraleitern. Beim Durchgang eines Schallstrahls durch einen Festkörper wird dessen Intensität durch Absorptionsvorgänge verringert. Im Leiter erfolgt das zu einem beträchtlichen Anteil durch die Phonon-Elektron-Wechselwirkung. Da in einem Supraleiter ein Teil der speziell für die Aufnahme von Energie geeigneten Elektronen an der Fermifläche gepaart auftritt, verringert sich beim Übergang NL → SL die Schallabsorption. Voraussetzung ist dabei – das ist aber experimentell ohne weiteres gesichert –, daß die Phononenenergie $\hbar\Omega < 2\Delta$ bleibt. Sonst würden durch die Elektron-Phonon-Wechselwirkung die Elektronenpaare aufbrechen.

Auch bei der Absorption elektromagnetischer Strahlung ist ein deutlicher Einfluß durch Elektronenpaare festzustellen. Derartige Versuche werden in mit Supraleitern beschichteten Hohlleitern ausgeführt. Wird die Frequenz der elektromagnetischen Strahlung verringert, dann tritt bei $h\nu = 2\Delta$ eine starke Abnahme der Absorption im Vergleich zum normalleitenden Zustand auf. Hier muß also zwischen Gebieten $h\nu < 2\Delta$ und $h\nu > 2\Delta$ unterschieden werden.

Tab. 8.5.2-8 Energielücken 2Δ in meV ausgewählter Supraleiter

Supraleiter	Zn	Cd	Al, V Ga	Sn, In, Tl, Ta	Nb	Pb	Hg
2Δ	2,4	3,2	3,3	3,4 3,5	3,6	4,3	4,6

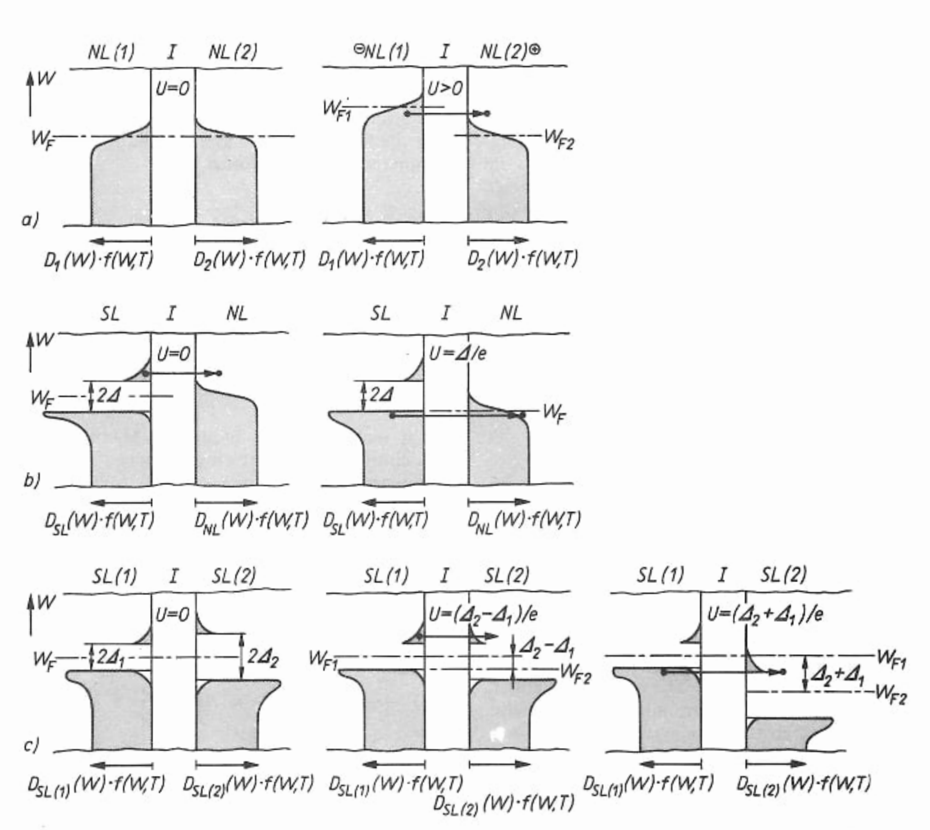


Abb. 8.5.2-7 Zustandsdichte und Einelektronentunnelung bei verschiedenen Grundstrukturen; der dominierende Tunnelungseffekt ist jeweils durch einen *Pfeil* markiert; a) Metall – Isolator – Metall, b) Supraleiter – Isolator – Metall, c) Supraleiter – Isolator – Supraleiter

Im ersten Fall reicht die eingestrahlte Quantenenergie nicht zum Aufbrechen der Elektronenpaare aus, während mit höherenergetischen Quanten die Bindung der Elektronenpaare aufgebrochen wird. Das ist dann aber der dem Normalleiter entsprechende Zustand. In beiden soeben diskutierten Beispielen erfolgt die Versuchsdurchführung des Vergleichs von NL und SL, indem man unterhalb T_K mit einem geeigneten Magnetfeld den Übergang NL → SL isotherm schaltet.

Der *Elektronenanteil der spezifischen Wärmekapazität* (↑ 8.4.3.5.) steigt bei Supraleitern unterhalb der Sprungtemperatur T_K exponentiell und nicht wie bei Normalleitern $\sim T$ an. Diese Temperaturabhängigkeit kann mit

$$(4) \quad c_{SL,e} \sim \exp(-\text{ konst. } \Delta/k_B T)$$

beschrieben werden. Auch hier tritt wieder deutlich die Wirkung der Energielücke im Anregungsspektrum zutage. Da es sich aber bei einem NL → SL-Übergang um einen Übergang 2. Ordnung handelt, tritt keine Umwandlungswärme auf.

Da unterhalb $T = T_K$ auch weniger ungepaarte Elektronen für die Wärmeleitung zur Verfügung stehen, muß man erwarten, daß ein Festkörper im supraleitenden Zustand entsprechend weniger Wärmeenergie transportieren kann als im normalleitenden. Das ist, wie Abb. 8.5.2-9 zeigt, tatsächlich der Fall. Größenordnungsmäßig gilt für das Verhältnis der Wärmeleitfähigkeiten $\lambda_{SL}/\lambda_{NL} \approx 1/10$. Diese Eigenschaft kann zur Konstruktion von *Wärmestromschaltern* ausgenutzt werden, indem man eine Metallbrücke zwischen zwei auf verschiedener Temperatur befindlichen Festkörper, z. B. in einem Kryostaten, durch ein überkritisches Magnetfeld schaltet.

8.5.3. Beschreibung der Supraleitung

8.5.3.1. Londontheorie. Von H. und F. LONDON sowie von M. VON LAUE wurde eine elektrodynamische Theorie der Supraleitung entwickelt, die von den elektromagnetischen Eigenschaften der Supraleiter ausgeht. Sie stellt eine Weiterentwicklung der Theorie von BECHER, SAUTER und HELLER dar, die einen Idealleiter zugrunde legten.

Ein *Idealleiter* hat die Eigenschaft $\varrho = 0$ bzw. $\sigma = \infty$. Mit der Annahme eines Idealleiters konnte aber der Meißner-Ochsenfeld-Effekt nicht erklärt werden. Ausgehend von Gleichung 8.4.(9) soll e durch q ersetzt werden, da ja zunächst nicht bekannt ist, welche Ladungsträger einzusetzen sind. Es gilt $mj/nq^2 = Et$ oder

$$(5) \quad \frac{\partial}{\partial t}\left(\frac{mj}{n_{SL}q_{SL}^2}\right) = E.$$

Gleichung 8.5.(5) wird als *1. Londongleichung* bezeichnet. Als Postulat muß dann gefordert

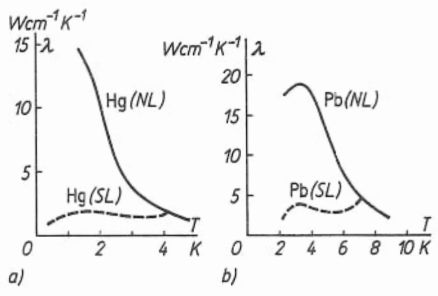


Abb. 8.5.2-9 Wärmeleitfähigkeit λ von Supraleitern; a) Quecksilber, b) Blei

werden, daß die Stromdichte im Supraleiter der Beziehung

$$(6) \quad j_{SL} = -\frac{n_{SL}q_{SL}^2}{m} A$$

genügt. A ist das Vektorpotential, das bekanntlich den folgenden Zusammenhang mit der magnetischen Induktion B

$$(7) \quad B = \frac{\partial}{\partial r} \times A$$

besitzt. Aus 8.5.(6) und 8.5.(7) ergibt sich die *2. Londongleichung*:

$$(8) \quad \frac{\partial}{\partial r} \times j_{SL} = -\frac{n_{SL}\, q_{SL}^2}{m}\left(\frac{\partial}{\partial r} \times A\right) = -\frac{n_{SL}\, q_{SL}^2}{m} B.$$

Zieht man an dieser Stelle die Maxwellschen Gleichungen mit verschwindender Verschiebungsstromdichte

$$(9) \quad \frac{1}{\mu_0}\left(\frac{\partial}{\partial r} \times B\right) = j_{SL}$$

und

$$(10) \quad \frac{\partial}{\partial r} B = 0$$

zu Rate, dann folgt

$$(11) \quad \frac{1}{\mu_0}\left(\frac{\partial}{\partial r} \times \frac{\partial}{\partial r} \times B\right) = \frac{\partial}{\partial r} \times j_{SL} = -\frac{n_{SL}q_{SL}^2}{m} B.$$

Das ist eine Bestimmungsgleichung für die Induktion B in einem Supraleiter. Mit Hilfe von

$$\frac{\partial}{\partial r} \times \frac{\partial}{\partial r} \times B = -\frac{\partial^2}{\partial r^2} B + \frac{\partial}{\partial r}\left(\frac{\partial}{\partial r} B\right)$$

findet man schließlich:

(12) $$\frac{\partial^2}{\partial r^2} B = \frac{n_{SL} q_{SL}^2 \mu_0}{m} B.$$

Gleichung 8.5.(12) hat die richtige Lösung für das Innere eines Supraleiters. Denn wenn $\boldsymbol{B}$ konstant sein soll, kann nur $\boldsymbol{B} = 0$ gelten. Das ist aber das Ergebnis des *Meißner-Ochsenfeld-Effekts*. Lassen wir nun in einer Randzone an der Oberfläche des Supraleiters ein Abklingen des Magnetfelds B von der Oberfläche $z = 0$ bis ins Innere z zu, dann hat 8.5.(12) die einfache Lösung

(13) $B_z = B_z(0) \exp(-z/\lambda_0)$.

Die Eindringtiefe λ_0 folgt zu $\lambda_0 = \left(\frac{m_{SL}}{\mu_0 n_{SL} q_{SL}^2}\right)^{1/2}$ oder, wenn wir im Vorgriff auf die Betrachtung der Theorie von Cooper $q_{SL} = 2e$, $m_{SL} = 2m_0$ und $n_{SL} = n/2$ setzen:

(14) $\lambda_0 = (m_0/\mu_0 n e^2)^{1/2}$.

Einsetzen der Elektronenmasse m_0, der Elektronenkonzentration n, der Permeabilität des Vakuums μ_0 und der Elementarladung e liefert z. B. für Zinn, wenn wir wieder annehmen, daß pro Zinnatom ein Leitungselektron bereitgestellt wird, $\lambda_0 = 28$ nm.

Das ist verglichen mit dem Meßwert $\lambda_{0Sn} = 51$ nm ein überraschend gutes Ergebnis. Weitergehende Schlüsse zu ziehen ist sicher nicht zulässig, da man annehmen muß, daß weder die Benutzung von m_0 noch von n in dieser Näherung exakt ist.

8.5.3.2. BCS-Theorie. Eine atomistische Theorie der Supraleitung wurde durch Bardeen, Cooper und Schrieffer 1957 entwickelt. Ausgangspunkt dieser Theorie ist das Vorhandensein von *Cooperpaaren*. Fröhlich und Bardeen konnten unabhängig voneinander zeigen, daß Elektronen im Gitter eines Festkörpers neben der abstoßenden Coulombwechselwirkung auch eine schwache Anziehungskraft aufeinander ausüben. Diese Anziehungskraft kommt dadurch zustande, daß jedes Leitungselektron die nächste Umgebung des Gitters polarisiert. Andere Elektronen, die solchen polarisierten Stellen im Gitter mehr als 1 μm nahekommen, erleben diese Gitterstellen als Potentialmulde und werden etwas angezogen.

Eine wirklich längerwährende Paarbildung erfahren nur Elektronen mit symmetrischer Lage im k-Raum, d. h. mit entgegengesetzt gleich großem Quasiimpuls $\hbar\boldsymbol{k}_1 = -\hbar\boldsymbol{k}_2$ oder $\boldsymbol{k}_1 + \boldsymbol{k}_2 = \boldsymbol{k}_{ges} = 0$.

Die Cooperpaare besitzen nun eine weitere wichtige Eigenschaft. Sie tragen infolge der Antiparallelstellung der Elektronenspins den Gesamtspin $\boldsymbol{S}_g = 0$. Als Bosonen haben sie die Tendenz, einen einzigen niedrigen Energiezustand zu besetzen. Es entsteht praktisch ein *makroskopischer Quantenzustand*.

Im wesentlichen können aber nur Elektronen in der Nähe der Oberfläche des Fermikörpers an dieser Bildung von Cooperpaaren teilnehmen. Die sich in der Energielücke dokumentierende Bindungsenergie der Cooperpaare, die auf eine Elektron-Phonon-Wechselwirkung zurückgeht, liegt in der Größenordnung von 5 meV. Es wird also nur eine dünne Schale des Fermikörpers in diese Wechselwirkung einbezogen. Eine einfache Abschätzung zeigt, daß nur ein kleiner Bruchteil der Leitungselektronen im zeitlichen Mittel zu Cooperpaaren geworden ist. Mit 8.4.(17) gilt für das Volumen des Fermikörpers im Impulsraum, wobei wir der Einfachheit halber eine Fermikugel annehmen:

(15) $(4/3)\,\pi p_F^3 = (4/3)\,\pi(2mE_F)^{3/2} = nh^3$.

Es gilt wieder, daß p_F der maximale Impuls, E_F die Fermienergie und n die Elektronenkonzentration sind, daraus folgt sofort

(16) $$\ln n = \frac{3}{2} \ln E_F + \ln \frac{4\pi(2m)^{3/2}}{3h^3}$$

und

(17) $$\frac{dn}{n} = \frac{3}{2} \frac{dE}{E_F}.$$

Setzen wir in 8.5.(17) für dE die Bindungsenergie der Cooperpaare ≈ 5 meV und den weiter oben abgeleiteten Wert von $E_F \approx 7$ eV für Leiter ein, dann erhalten wir die wichtige Aussage, daß das Verhältnis der Konzentration von gepaarten Elektronen $n_s \mathrel{\hat{=}} dn$ zur Gesamtkonzentration der Elektronen n $n_s/n \approx 10^{-4}$ ist.

Aus dieser Betrachtung folgt aber auch sofort, daß die kinetische Energie der Cooperpaare $E_{k,c} \approx 2E_F - \Delta$ sein muß. Darum liegt die Gesamtenergie des Supraleiters unterhalb der Gesamtenergie des Normalleiters bei gleicher Temperatur. Anders betrachtet, um den supraleitenden Zustand in den normalleitenden zu überführen, wird der Energiebetrag $n_s\Delta$ zum Aufbrechen der Bindung der Cooperpaare benötigt. Der Unterschied zwischen Leitungselektronen und Cooperpaaren wird auch beim Stromtransport unter der Wirkung eines elektrischen Felds deutlich. Zu dem Quasiimpuls eines Leitungselektrons im feldfreien Fall muß der durch die Elektronendrift bedingte Zusatzimpuls $d(\hbar k)$ addiert werden (↑ 8.4.4.3.). Diese Impulsänderung ist maximal von der Größenordnung $10^{-9}\,\hbar k_F$. Die Cooperpaare besitzen im stromlosen Fall einen Schwerpunktimpuls Null. Bei Stromfluß ist der Gesamtimpuls der Cooperpaare nur ganz wenig von Null verschieden. Seinen möglichen Maximalwert findet man durch die Überlegung, daß es für jeden Supraleiter eine kritische Stromdichte gibt, die

nicht überschritten werden darf, weil dann die Cooperpaare aufbrechen. Maximal können also Cooperpaare den Schwerpunktimpuls $\hbar k_{S,max} \approx (4\Delta m)^{1/2} \approx 0{,}027\hbar k_F$ besitzen. Da Bosonen bevorzugt einen einzigen, den niedrigstmöglichen Energiezustand besetzen, besitzen folgerichtig alle Cooperpaare dieselbe Energie und denselben Impuls sowie in Anwendung des Welle-Teilchen-Dualismus die gleiche Frequenz und die gleiche Wellenlänge.

Aus der Quantenmechanik ist bekannt, daß die Phase eines Zustands um so genauer angegeben werden kann, je stärker dieser Zustand besetzt ist. Das führt aber zu der wichtigen Aussage, daß in einem Supraleiter die Materiewellen aller Cooperpaare in gleicher Phase schwingen. Die Cooperpaare können alle durch eine einzige Wellenfunktion ψ_c beschrieben werden. Es gilt dann: $n_s/2 \approx \psi_c\psi_c^*$, d. h., das Amplitudenquadrat der Wellenfunktion der Cooperpaare bestimmt die Konzentration der Cooperpaare $n_s/2$. Die Größen n_s und $|\psi_c|$ steigen von Null ($T = T_k$) stetig auf ihren Maximalwert bei $T = 0$ K an.

8.5.4. Josephsoneffekte, Anwendungen

8.5.4.1. Technische Anwendung der Supraleiter. Gegenwärtig werden Supraleiter 3. Art soweit technologisch beherrscht, daß man aus ihnen zuverlässige Supraleiterkabel einheitlicher Qualität herstellen kann. Zwei volkswirtschaftlich besonders bedeutsame Randbedingungen scheinen ihren technischen Einsatz ganz besonders zu fördern. Das sind die Einsparung von Elektroenergie und die Materialökonomie. Supraleiter haben deshalb überall dort große Einsatzchancen, wo es um die Weiterleitung großer Mengen von Elektroenergie und um die Erzeugung starker Magnetfelder geht. In einigen Forschungszentren laufen inzwischen Versuchsstrecken zur Lösung der Einkopplungs- und Auskopplungsprobleme an Supraleiter/Normalleiter-Kontaktstellen, der Probleme der Konstruktion *supraleitender Kabel* und der zugehörigen Kryotechnik. Mit Supraleiterwicklung können starke Magnetfelder weit ökonomischer erzeugt werden als mit konventionellen Methoden. Neben einer Einsparung an ferromagnetischen Materialien und einer enormen Gewichtseinsparung fällt ganz besonders die Vermeidung Joulescher Wärmeverluste beim Stromdurchgang ins Gewicht. Diese Vorteile werden bei kleineren Anlagen heute noch durch den hohen Kostenaufwand für die notwendige Kryotechnik unter Einsatz von flüssigem Helium zunichte gemacht. Bei Großanlagen ist aber die Anwendung von Supraleitern heute schon kostengünstiger. Eingesetzt werden supraleitende Magnete zur Zeit bis zu Abmessungen von 10 m³ Volumen. Zum Beispiel werden bestehende Teilchenbeschleuniger auf *supraleitende Ablenkmagnete* umgerüstet. Wo das schon geschehen ist, konnten die erreichbaren Teilchenenergien der verschiedensten Beschleunigertypen mindestens verdoppelt werden. Die jetzt im Bau befindlichen Großbeschleuniger einer neuen Generation mit Endenergien von 500 GeV bis 1 TeV werden von vornherein mit *Supraleitermagneten* ausgerüstet. Das gilt auch für moderne Teilchendetektoren, in denen Ablenkmagnete installiert werden müssen (Tafeln 46/47). Einen hohen Entwicklungsstand haben Supraleitermagnete für Fusionsreaktoren vom Tokamaktyp erreicht. Bekanntlich werden hier höchste Magnetfelder zur Einschnürung des hocherhitzten Plasmas benötigt. Auch zur Trennung von beschleunigten Ionen in MHD-Generatoren sind supraleitende Magnete schon vorteilhaft im Einsatz. Es gibt auch eine ganze Reihe von Teststrecken für *Magnetkissenzüge*, die mindestens Geschwindigkeiten von 500 kmh⁻¹ zulassen werden, bei denen supraleitende Spulen zumindest einen Teil der Stützfelder liefern. Auch supraleitende Energiespeicher sind denkbar. Natürlich haben supraleitende Magnete auch eine enorme Bedeutung für die Forschungsarbeiten, bei denen extrem hohe Magnetfelder benötigt werden. Interessanterweise werden die höchsten Magnetfeldstärken heute durch eine Überlagerung des Felds eines supraleitenden Magneten als Außenspule und des Felds einer im Inneren angeordneten Spule mit konventioneller Bauweise erreicht. Diese Kombination ist erforderlich, um die zur Zeit gegebene Grenze zu überschreiten, die durch die kritische Feldstärke H_{k2} von Supraleitern gegeben ist. Bisher wurden Spitzenwerte für die magnetische Induktion $B \approx 20T$ im Dauerbetrieb erreicht. Eine Weiterführung dieser Bemühungen, um noch höhere Induktionswerte zu erreichen, ist von hohem wissenschaftlichen Interesse, weil noch völlig offen ist, wie sich Stoffe unter extrem hohen Magnetfeldern verhalten werden.

8.5.4.2. Wesen der Josephsoneffekte. Josephson konnte 1962 aufgrund theoretischer Überlegungen eine Reihe von Experimenten an einem Übergang: Supraleiter 1 – Isolator – Supraleiter 2 vorschlagen. Schon 1963 wurden diese vorausgesagten Effekte durch verschiedene Forschungsgruppen im vollen Umfang realisiert.

Das Wesen der *Josephsoneffekte* besteht darin, daß hier das Verhalten zweier schwach gekoppelter *makroskopischer Quantensysteme* zugrunde liegt. Damit konnte man derartige Quantensysteme erstmalig umfassend untersuchen. Die schwache Kopplung von zwei Supraleitern wird erreicht, wenn man z. B. eine sehr dünne Isolierschicht zwischen diesen anordnet, die eine

Dicke von $\approx$ 2 nm hat. Physikalisch liegt allen Josephsoneffekten die *Tunnelung* von Cooperpaaren durch eine dünne Barriere zugrunde.
Diese Aussage ist natürlich äquivalent der Feststellung, daß sich die Wellenfunktionen von Supraleiter 1 und Supraleiter 2 schwach überlappen.
Für zwei räumlich getrennte Supraleiter 1 und 2 können die Wellenfunktionen durch

$$(18)\quad \begin{aligned} \psi_1(r, t) &= n_{SL1}^{1/2}(r,t) \exp[i\varkappa_1(r, t)] \\ \psi_2(r, t) &= n_{SL2}^{1/2}(r,t) \exp[i\varkappa_2(r, t)] \end{aligned}$$

beschrieben werden.
Da $\varkappa(r, t)$ die Phasenlage aller Cooperpaare eines Supraleiters beschreibt, wird damit ein makroskopischer Zustand festgelegt, und folgerichtig kann diese Phasenlage hier wirklich auch experimentell beobachtet werden.
Das ist insofern neu und eine glänzende Bestätigung theoretischen Denkens, weil die Quantentheorie die Phasenlage des einzelnen Quantenzustands seit vielen Jahrzehnten formuliert hatte, die Modelle dadurch auch völlig widerspruchsfrei wurden, aber bei den Beobachtungswerten, infolge bestimmter Mittelwertbildungen, Phasenlagen explizit keine Rolle spielten. Es ist nun leicht einzusehen, daß sich die Phasenlage $\varkappa(r,t)$ eines Supraleiters erst mit Herausbildung einer größeren Anzahl von Cooperpaaren einstellen wird. Das ist also unmittelbar beim Übergang Normalleiter/Supraleiter der Fall und vielen Zufälligkeiten unterworfen, so daß jeder Supraleiter eine für jeden Abkühlungsvorgang spezifische Phasenlage hat. Anders ausgedrückt heißt das, wenn wir zwei Supraleiter schwach koppeln, dann läßt sich feststellen, daß diese Kombination eine zufällige von Null verschiedene Phasendifferenz $\Delta\varkappa(1/2)$ aufweist. Derartige Kopplungsstellen sind in einigen Ausführungsformen in Abb. 8.5.4-1 angegeben. Hieraus ist ersichtlich, daß eine schwache Kopplung von Supraleitern durch eine dünne Zwischenschicht von nichtsupraleitendem Material, wie Isolierstoffen, bevorzugt Oxiden, Halbleitern und Normalleitern hindurch erfolgen kann. Es ist aber auch möglich, einfach eine Einschnürstelle im Supraleiter selbst anzubringen.

8.5.4.3. Gleichstrom-Josephsoneffekt. Ausgehend von einem einfachen Ansatz für die zeitabhängige Schrödingergleichung der Kopplungsstelle zwischen zwei Supraleitern

$$(19)\quad \begin{aligned} &\text{SL 1: } i\hbar \frac{\partial \psi_1}{\partial t} = \mu_1\psi_1 + W_T\psi_2, \\ &\text{SL 2: } i\hbar \frac{\partial \psi_2}{\partial t} = \mu_2\psi_2 + W_T\psi_1 \end{aligned}$$

kann der Stromfluß von Supraleiter zu Supraleiter berechnet werden. Rechts stehen in den Gleichungen die chemischen Potentiale μ_1 und μ_2 der Cooperpaare und W_T die Transferwechselwirkungsenergie, womit die schwache Kopplung $|W_T| \ll |\mu|$ berücksichtigt wird.
Setzt man 8.5.(18) in 8.5.(19) ein, dann ergibt das:

$$(20)\quad \begin{aligned} &\frac{i\hbar}{2} n_{SL1}^{-1/2} \frac{\partial n_{SL1}}{\partial t} \exp(i\varkappa_1) \\ &\quad - \hbar n_{SL1}^{1/2} \exp(i\varkappa_1) \frac{\partial \varkappa_1}{\partial t} \\ &= \mu_1 n_{SL1}^{1/2} \exp(i\varkappa_1) + W_T n_{SL2}^{1/2} \exp(i\varkappa_2) \end{aligned}$$

und

$$\begin{aligned} &\frac{i\hbar}{2} n_{SL2}^{-1/2} \frac{\partial n_{SL2}}{\partial t} \exp(i\varkappa_2) \\ &\quad - \hbar n_{SL2}^{1/2} \exp(i\varkappa_2) \frac{\partial \varkappa_2}{\partial t} \\ &= \mu_2 n_{SL2}^{1/2} \exp(i\varkappa_2) + W_T n_{SL1}^{1/2} \exp(i\varkappa_1). \end{aligned}$$

Berücksichtigt man nun, daß $\exp i(\varkappa_2 - \varkappa_1) = \cos(\varkappa_2 - \varkappa_1) + i \sin(\varkappa_2 - \varkappa_1)$ ist, dann lassen sich Real- und die Imaginärteile in 8.5.(20) trennen und man erhält

$$(21)\quad \begin{aligned} &\frac{\partial n_{SL1}}{\partial t} + 2i\, n_{SL1} \frac{\partial \varkappa_1}{\partial t} \\ &= -\frac{2i}{\hbar} [\mu_1 n_{SL1} + W_T(n_{SL1} n_{SL2})^{1/2} \\ &\quad \times \exp i(\varkappa_2 - \varkappa_1)] \end{aligned}$$

und

$$\begin{aligned} &\frac{\partial n_{SL2}}{\partial t} + 2i\, n_{SL2} \frac{\partial \varkappa_2}{\partial t} = -\frac{2i}{\hbar} \\ &\times [\mu_2 n_{SL2} + W_T(n_{SL1} n_{SL2})^{1/2} \exp i(\varkappa_1 - \varkappa_2)]. \end{aligned}$$

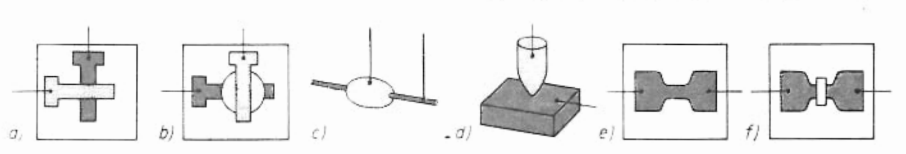


Abb. 8.5.4-1 Typen von Josephsonübergängen; a) gekreuzte Leiterbahnen (ein Supraleiter 1 nm dick oxidiert), b) mit isolierender oder halbleitender Zwischenschicht, c) oxidierter supraleitender Draht mit supraleitendem Löttropfen, d) Punktkontakt zwischen zwei Supraleitern, e) Einschnürung zwischen zwei Supraleitergebieten ($\approx 1 \times 2\ \mu m^2$), f) weitgehende Entkopplung von zwei Supraleitern an einer dünnen Brücke ($\approx 10 \times 10\ \mu m^2$) durch eine Metallschicht

Das führt dann sofort zu

$$(22)\quad \frac{\partial n_{SL1}}{\partial t} = \frac{2W_T}{\hbar}(n_{SL1}n_{SL2})^{1/2}\sin(\varkappa_2 - \varkappa_1)$$

und

$$\frac{\partial \varkappa_1}{\partial t} = -\frac{1}{\hbar}\left[\mu_1 + W_T\left(\frac{n_{SL1}}{n_{SL2}}\right)^{1/2} \times \cos(\varkappa_2 - \varkappa_1)\right].$$

Nimmt man nun zur Vereinfachung gleiche Supraleiter $n_{SL1} = n_{SL2}$ an und hält die Temperatur konstant, daß die Zahl der Cooperpaare gleich bleibt, $\partial n_{SL1}/\partial t = \partial n_{SL2}/\partial t$, dann vereinfacht sich Gleichung 8.5.(22) zu

$$(23)\quad \frac{\partial n_{SL}}{\partial t} = \frac{2W_T}{\hbar} n_{SL}\sin\Delta\varkappa(1,2);$$

mit $\Delta\varkappa(1,2) = \varkappa_2 - \varkappa_1$.

Da die Stromstärke die Kopplungsstelle I_S der Beziehung $I_S \sim \frac{\partial n_{SL}}{\partial t}$ genügen muß, erhält man für den Tunnelstrom I_S, ohne daß ein äußeres Feld anliegt,

$$(24)\quad I_S = I_{S,max}\sin\Delta\varkappa.$$

Dieses erstaunliche Resultat besagt:
In einem schwach gekoppelten System zweier Supraleiter fließt ein Strom, der einen je nach Phasenlage der beiden Supraleiter zufälligen Wert zwischen $\pm I_{S,max}$ hat. Der Strom $I_{S,max}$ hat den Wert $I_{S,max} = 4W_T e_0\, n_{SL}V/\hbar$.

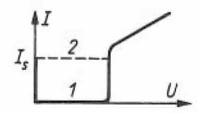


Abb. 8.5.4-2 Strom-Spannungs-Kennlinie eines Josephsontunneleffekts

Dabei sind e_0 Elementenladung und V das Volumen der supraleitenden Anordnung. Auf der Grundlage dieser Überlegungen wird es verständlich, warum die Kennlinie des Tunnelstroms zwischen schwach gekoppelten Supraleitern den in Abb. 8.5.4-2 skizzierten Verlauf besitzt. Reale Anordnungen können zwischen den beiden metastabilen Kennlinienästen 1 und 2 hin und her geschaltet werden. Damit ist ein *Josephsontunnelelement* als Speichereinheit geeignet.

8.5.4.4. Wechselstrom-Josephsoneffekt. Ausgehend von den bisher durchgeführten Rechnungen kann man aus 8.5.(22) die beiden Gleichungen:

$$(25)\quad \frac{\partial \varkappa_1}{\partial t} = -\frac{1}{\hbar}[\mu_1 + W_T\cos(\varkappa_2 - \varkappa_1)]$$

und

$$\frac{\partial \varkappa_2}{\partial t} = -\frac{1}{\hbar}[\mu_2 + W_T\cos(\varkappa_1 - \varkappa_2)]$$

erhalten. Sucht man jetzt $\partial(\Delta\varkappa/\partial t)$ auf, so findet man:

$$(26)\quad \frac{\partial(\Delta\varkappa)}{\partial t} = \frac{\partial\varkappa_2}{\partial t} - \frac{\partial\varkappa_1}{\partial t} = \frac{1}{\hbar}(\mu_1 - \mu_2).$$

Wie schon ausgeführt, ist es durch Anlegen eines äußeren Felds möglich, Cooperpaaren eine zusätzliche Energie zuzuführen. Legt man an die beschriebenen Anordnungen eine äußere Spannung U an, dann verschiebt sich das chemische Potential der makroskopischen Quantensysteme 1 und 2 gegeneinander wie

$$(27)\quad \mu_1 - \mu_2 = 2e_0U.$$

Das heißt aber $\frac{\partial(\Delta\varkappa)}{\partial t} = \frac{2e_0U}{\hbar}$ und nach Integration $\Delta\varkappa = \Delta\varkappa_0 + \frac{2e_0U}{\hbar}t = \Delta\varkappa_0 + 2\pi\nu t$.

Setzt man dieses Ergebnis in 8.5.(24) ein, dann ergibt sich

$$(28)\quad I_S = I_{S,max}\sin\{(\Delta\varkappa_0) + 2\pi\nu t\}.$$

Diese Gleichung läßt sich so interpretieren, daß durch Anlegen einer Spannung die Phasenlage der Supraleiter zueinander verändert wird und ein Wechselstrom durch die Anordnung fließt. Dieser Wechselstrom hat die Frequenz:

$$(29)\quad \nu = 2e_0U/\hbar.$$

Liegt z. B. 1 μV an, wird ν = 484 MHz. Ein Josephsonelement strahlt also eine Mikrowellenstrahlung ab. Gleichung 8.5.(29) ist nichts anderes als der Energiesatz

$$(30)\quad h\nu = 2e_0U$$

und Ausdruck der Umsetzung von potentieller Energie eines Potentialgefälles U in kinetische Energie der Cooperpaare $2e_0U$. Dieser Energie-

Tab. 8.5.4-3 Bestimmung von Naturkonstanten mit Hilfe des Josephsoneffekts nach CLARK

Elementarladung	e in 10^{-19} As	1,6021917	± 4,4 ppm
Plancksches Wirkungsquantum	h in 10^{-34} Ws²	6,626196	± 7,6 ppm
Elektronenmasse	m in 10^{-31} kg	9,109558	± 6,0 ppm
Loschmidtzahl	$L \cdot 10^{26}$	6,022169	± 6,6 ppm
Feinstrukturkonstante	$\alpha - 1$	137,03602	± 1,5 ppm

betrag wird abgestrahlt, wenn die Cooperpaare von ihrem Ausgangsquantenzustand μ_1 auf den Endzustand μ_2 zurückfallen. Auch dieser Effekt wird experimentell exakt bestätigt. Exakt heißt in diesem Falle, daß offensichtlich der Effekt genauer arbeitet, als es die Kenntnis der Naturkonstanten e_0 und $\hbar$ sowie die Definition der Spannungseinheit Volt nachzuprüfen gestattet. Man konnte aber inzwischen mit Hilfe dieses Effekts die Naturkonstanten genauer bestimmen (Tab. 8.5.4-3). Schließlich sei erwähnt, daß durch Einstrahlung einer Mikrowelle in einen Josephsonübergang auch Beziehung 8.5.(30) in umgekehrter Reihenfolge funktioniert, d. h., es entsteht eine entsprechende Spannung am Übergang.

8.5.4.5. Josephsonübergänge im Magnetfeld. Aus zwei Josephsonübergängen der in Abb. 8.5.4-1 skizzierten Typen kann ein supraleitendes *Quanten-Interferenz-Bauelement* (*SQID*) aufgebaut werden (Abb. 8.5.4-4). Es besteht also aus zwei parallel geschalteten Übergängen, die eine kleine Fläche A umschließen, oder es liegt ein supraleitender Ring vor, der hier allerdings aus zwei nur schwach gekoppelten Supraleitern besteht. Die Wirkungsweise einer derartigen Anordnung läßt sich aufklären, wenn man zunächst berechnet, welche Phasendifferenz $\Delta\varkappa$ sich einstellt, wenn die Fläche A einen magnetischen Fluß Φ einschließt. Bekanntlich muß bei Betrachtung der Bewegung von Ladungen q in einem elektromagnetischen Feld der kanonisch konjugierte Impuls als $p = mv + qA$ angesetzt werden. A ist hier wieder das Vektorpotential. Im Falle von Cooperpaaren lautet diese Beziehung

(31) $p_c = 2m_0v_c - 2eA.$

Da nun der Raum, in dem sich die Cooperpaare bewegen, selbst aufgrund des Meißner-Ochsenfeld-Effekts völlig frei vom Mangetfeld bleibt, muß, da auf die Cooperpaare keine Kraftwirkung ausgeübt wird, $|2m_0v_c| =$ konst. gelten. An dieser Stelle sei aber darauf hingewiesen, daß an den Kopplungsstellen aufgrund der endlichen Eindringtiefe des Magnetfelds ein Eindringen des magnetischen Flusses in den supraleitenden Ring möglich ist. Das wäre bei starker Kopplung nicht der Fall. Wenn die durch ein Magnetfeld entstehenden Phasendifferenzen berechnet werden sollen, muß für den einzelnen Josephsonkontakt von Abb. 8.5.4-4b ausgegangen werden. Hier sind die Gebiete hervorgehoben, in die ein äußeres Magnetfeld einzudringen vermag. Eine irgendwie geartete Zwischenschicht wird ebenso durchdrungen wie ein der Eindringtiefe entsprechendes Gebiet an der Oberfläche der Supraleiter. Man kann nun unter Beachtung von $|p| = \dfrac{h}{\lambda'}$ folgendes ansetzen (λ' ist hier die Wellenlänge der Materiewelle)

(32) $$\Delta\varkappa_B = 2\pi \int \frac{ds}{\lambda} = \frac{1}{\hbar} \int p_c \, ds,$$
$$\Delta\varkappa_B = -\frac{2e}{\hbar} \int A \, ds$$

Die durch das Magnetfeld bedingte maximale Phasendifferenz $\Delta\varkappa_B$ ergibt sich mit Hilfe des Stokesschen Satzes und der Definition 8.5.(7)

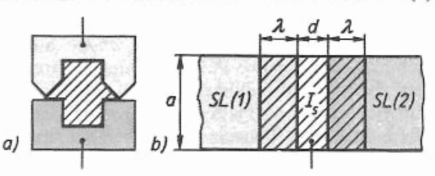


Abb. 8.5.4-4 a) Prinzipieller Aufbau eines Quanten-Interferenz-Bauelements, b) von einem äußeren Magnetfeld beeinflußter Bereich an einem Josephsonübergang der Isolatordicke d

auf einem geschlossenen Weg 1 – 2 – 3 – 4 – 1 um den Josephsonübergang (↑ Abb. 8.5.4-4b)

$$\Delta\varkappa_B = -\frac{2e}{\hbar} \oint A \, ds = -\frac{2e}{\hbar} \int \left(\frac{\partial}{\partial r} \times A\right) df;$$

(df ist das Flächenelement)

(33) $$\Delta\varkappa_B = -\frac{2e}{\hbar} \int B \, df = -\frac{2e}{\hbar} \Phi$$
$$\Phi = B(2\lambda + d)\, a.$$

Da natürlich auch ohne Magnetfeld eine Phasendifferenz $\Delta\varkappa$ zwischen den Supraleitern 1 und 2 vorhanden ist, kann die Phasenlage der beiden Supraleiter durch

(34) $$\Delta\varkappa_1 = \Delta\varkappa + \Delta\varkappa_B/2$$
$$\Delta\varkappa_2 = \Delta\varkappa - \Delta\varkappa_B/2$$

beschrieben werden. Analog 8.5.(24) wird der maximale Strom durch beide Josephsonübergänge unter der Wirkung eines Magnetfelds

(35) $$I_{sB,max} = I_{s,max} \{\sin(\Delta\varkappa + \Delta\varkappa_B/2) + \sin(\Delta\varkappa - \Delta\varkappa_B/2)\}$$
$$I_{sB,max} = I_{s,max} \sin \Delta\varkappa \cos(e\Phi/\hbar).$$

Für das SQID in Abb. 8.5.4-4a muß der Integrationsweg um die Fläche A herumgeführt werden. Das führt auf $\Phi = BA$, falls B die Fläche A senkrecht durchstößt. Die Funktion 8.5.(35) hat Maxima bei $\Phi = n\Phi_0 = nh/2e$ und Minima bei $\Phi = (2n + 1)\,\Phi_0/2$. Φ_0 ist das Flußquant, n eine ganze Zahl.

8.5.4.6. Anwendung von Josephsonelementen. Da mit Hilfe eines SQID's an Hand einer Strommessung noch sehr gut Magnetfeldänderungen von der Größenordnung eines *Flußquants* $\approx 2 \cdot 10^{-15}$ Wb erfaßt werden können, stellt man auf dieser Basis sehr genaugehende Magnetometer her. Bei einer wirksamen Fläche $A = 2\,\text{cm}^2$ können noch Änderungen der magnetischen Induktion $\Delta B \approx 10^{-12}$ T nachgewiesen

werden. Da sich sehr leicht Anordnungen herstellen lassen, die es gestatten, ein Magnetfeld als stromproportionale oder spannungsproportionale Größe zu erzeugen, sind Josephsonelemente auch als sehr empfindliche Galvanometer verwendbar. Inzwischen werden *supraleitende Magnetometer* gefertigt und für die geologische Erkundung eingesetzt. Mit Hilfe derartiger Instrumente war es möglich, *Magnetokardiogramme* und *Magnetoenzephalogramme* aufzunehmen. Die bei der Paartunnelung auftretenden Kennlinien bildeten die Basis für die Entwicklung von *Kryospeichern.* Josephsonspeichereinheiten stellen einen weiteren Schritt zur Miniaturisierung von Bauelementen dar. Hier werden derzeit Spitzenwerte der bit-Dichte pro Flächeneinheit erreicht. Auch aus dieser Sicht wird das Interesse an Supraleitern mit Sprungtemperaturen oberhalb Zimmertemperatur begründet. Eine mögliche Einführung einer *Josephsonlogik* in die Rechentechnik kann nicht ausgeschlossen werden.

8.6. Dielektrische und ferroelektrische Erscheinungen in Festkörpern

8.6.1. Typische Eigenschaften der Dielektrika

8.6.1.1. Bedeutung der Dielektrika und Isolierstoffe. In der Elektrotechnik und Elektronik begegnen uns Isolierstoffe der verschiedensten Art und für eine Reihe ganz wesentlicher Anwendungsgebiete. Insbesondere durch die stürmische Entwicklung der Hochpolymere, aber auch durch die Weiterentwicklung keramischer, glasartiger und oxidischer Isolierstoffe gelang es, die Zuverlässigkeit und den Gebrauchswert elektrotechnischer und elektronischer Erzeugnisse wesentlich zu erhöhen. Mit der immer besseren Beherrschung der Isolationsprobleme an Freileitungsmasten und in Umspannanlagen konnten riesige Verbundnetze zur Fernübertragung von Elektroenergie über Ländergrenzen hinweg installiert werden, wobei heute Hochspannungsgleichstromübertragungen von 500 bis 700 kV benötigt werden, um die Energieverluste möglichst klein zu halten. Ein besonderes Problem ist die immer weiter fortschreitende Miniaturisierung elektronischer Bauelemente. Diese Entwicklung geht zwar mit einer Absenkung der Betriebsspannungen einher, erfordert aber trotzdem, daß die galvanische Trennung bei Feldstärken von 1 $MVcm^{-1}$, insbesondere bei der Gateisolation beherrscht wird. In der Fertigung mikroelektronischer Schaltkreise ist natürlich auch die Abscheidung bestimmter Isolatordicken, die Strukturierung von Isolatorschichten und ihre Alterungsbeständigkeit entscheidend. Auch die gesamte Meßtechnik hat von der Weiterentwicklung der Dielektrika außerordentlich profitiert. Heute können sehr kleine Ströme sicher gemessen werden, weil es durch die Entwicklung hochohmiger Isolierstoffe gelang, die sonst störenden Leck- und Kriechströme immer weiter herabzudrücken. Das Verhältnis von räumlicher Abmessung und Kapazität konnte bei Kondensatoren immer weiter gesenkt werden, weil auch wesentliche Fortschritte bei den Kondensatordielektrika erreicht wurden.

8.6.1.2. Eigenschaften der Dielektrika. Auch auf dem Gebiet der Isolierstoffe muß man heute davon ausgehen, daß noch viele Zusammenhänge zwischen dem mikroskopischen und makroskopischen Bau und den Eigenschaften besser geklärt werden müssen. Die typischen Eigenschaften eines Dielektrikums lassen sich aber so fassen, daß ein klares Bild dieser Stoffklasse entsteht.

Bringt man ein Dielektrikum zwischen die Platten eines Kondensators, dann erhöht sich dessen Kapazität gegenüber dem Vakuumwert. Die elektrische Leitfähigkeit eines Dielektrikums ist bei Zimmertemperatur sehr klein. Das liegt daran, daß in ihnen nur wenige oder praktisch keine freien Ladungsträger zur Verfügung stehen. Außerdem besitzen Dielektrika sehr kleine Ladungsträgerbeweglichkeiten. Man kann aber durch verschiedene Maßnahmen erreichen, daß sich die Leitfähigkeit stark erhöht. Das ist auf jeden Fall immer durch Temperaturerhöhung möglich, aber oft auch durch Anlegen hoher elektrischer Feldstärken $> 1\ MVcm^{-1}$ unter Zuhilfenahme von Metall- oder Elektrolytelektroden, die es gestatten, große Ladungsträgermengen zu injizieren. Bei hohen elektrischen Feldstärken werden Dielektrika irreversibel zerstört, wobei sich sowohl chemische als auch strukturelle Veränderungen vollziehen. Diesen Vorgang bezeichnet man als elektrischen Durchschlag. Dielektrika sind meist sehr schlechte Wärmeleiter.

Von ihrem Aufbau her sind Dielektrika vorwiegend durch heteropolare oder durch Van-der Waals-Bindungskräfte zwischen ihren atomaren und molekularen Bausteinen gekennzeichnet.

8.6.1.3. Geeignete Stoffe für Dielektrika. Aufgrund der oben beschriebenen Merkmale umfassen die Dielektrika eine Fülle von Stoffen und Stoffgruppen, die hier in der Ordnung ihrer zunehmenden Komplexität zusammengestellt werden sollen. Zunächst müssen die atomaren Elemente Helium, Neon, Argon, Krypton, Xenon und Radon genannt werden. Auch die molekularen Elemente Stickstoff, Sauerstoff, Fluor, Chlor und Brom gehören dazu. Viele anorganische Stoffe wie die meisten Salze, z. B. Fluoride, Chloride, Bromide, Iodide, Sulfate und andere, die meisten Gläser, keramische Werkstoffe, viele Minerale, insbesondere Glim-

mer und Asbest, und die für die Elektronik so wichtigen Oxide wie SiO_2, SiO, Al_2O_3, Ta_2O_5 und Nitride wie Si_3N_4, Ge_3N_4 sind Dielektrika. Das gilt auch für Wasser und Eis in reinem Zustand.

Einen sich noch ständig erweiternden Vorrat von dielektrischen Stoffen stellt die organische Chemie bereit. Dazu muß man nahezu alle (>99%) monomolekularen Verbindungen, wie Kohlenwasserstoffe, Heterocyklen, metallorganische Verbindungen, Komplexsalze und andere rechnen und vor allem die vielen auch volkswirtschaftlich sehr bedeutsamen polymeren Werkstoffe wie Polystyren, Polyethylen, Polyvinylchlorid, Polyamide, Polytetrafluorkarbone, Polyester, Polyurethane, Parylene, Epoxidharze, Polyterephthalat, Polysiloxane.

8.6.2. Beschreibung der dielektrischen Erscheinungen

8.6.2.1. Dielektrizitätskonstante und Verlustwinkel. Die Kapazität des Kondensators beträgt im Vakuum

(1) $C_0 = Q_0/U$ und (2) $C = Q/U$,

wenn sich ein Dielektrikum im Kondensator befindet. Das Verhältnis der Kapazitäten bzw. Ladungsmengen $C/C_0 = Q/Q_0 = \varepsilon$ ist die *relative Dielektrizitätskonstante*. Für die meisten in der Praxis auftretenden Fälle kann ε wie eine skalare Größe behandelt werden. Für anisotrope Einkristalle muß man allerdings berücksichtigen, daß die relative Dielektrizitätskonstante ein Tensor ist. Im allgemeinen Fall lautet der Zusammenhang zwischen Verschiebung $\boldsymbol{D}$ und elektrischer Feldstärke $\boldsymbol{E}$

(3) $\boldsymbol{D} = \varepsilon\varepsilon_0 \boldsymbol{E}$

und für die *Polarisation* $\boldsymbol{P}$ ergibt sich in jedem Volumenelement des Dielektrikums

(4) $\boldsymbol{P} = \boldsymbol{D} - \varepsilon_0 \boldsymbol{E} = (\varepsilon - 1)\, \varepsilon_0 \boldsymbol{E} = \varkappa_{el} \varepsilon_0 \boldsymbol{E}$.

Die Größe $\varkappa_{el}$ wird elektrische Suszeptibilität genannt. Setzt man ε als rein reelle Größe an, dann hat man noch nicht berücksichtigt, daß sich ein Dielektrikum in einem Wechselfeld immer, wenn auch sehr schwach, erwärmt. Der Vorgang der Polarisation erfolgt also nicht reibungsfrei und die Phase der zeitlich veränderlichen Größe Polarisation hängt der Phase des Wechselfelds nach.

Dies wird durch eine komplexe relative Dielektrizitätskonstante $\varepsilon = \varepsilon' - \varepsilon''$ beschrieben.

Der Verlustwinkel errechnet sich mit dieser Festlegung zu $\tan \delta = \varepsilon''/\varepsilon'$.

8.6.2.2. Polarisation eines Dielektrikums. Eine genauere Betrachtung des Vorgangs der Polarisation gewährt einen tieferen Einblick in die mikroskopisch ablaufenden Prozesse in einem Dielektrikum, an das ein zeitlich veränderliches elektrisches Feld angelegt wurde. Beschreibt man Gleichung 8.6.(4) verbal, dann läßt sich die Polarisation als der Anteil der dielektrischen Verschiebung verstehen, der durch die Stoffeigenschaften des Dielektrikums bestimmt wird. Die Polarisation $\boldsymbol{P}$ kann näherungsweise als die Summe aller elektrischen Dipolmomente $\boldsymbol{p}$ pro Volumeneinheit berechnet werden:

(5) $\boldsymbol{P} = \sum \boldsymbol{p}_i$.

Derartige Dipole können als *permanente Dipole* im Dielektrikum immer vorhanden sein bzw. erst durch die Einwirkung eines elektrischen Felds entstehen. Im zweiten Fall nennt man sie *induzierte Dipole*.

Bei möglichst vollkommener Ausrichtung der permanenten Dipole in einem elektrischen Feld wird von *Orientierungspolarisation* gesprochen. Dielektrika, die vorzugsweise diesem Polarisationsprozeß unterliegen, also genügend permanente Dipole haben, bezeichnet man als *Parelektrika*. Da ein elektrischer Dipol als minimale Einheit zwei gleichgroße Ladungen q verschiedener Polarität besitzt, die sich in einem endlichen Abstand s, $p = qs$ befinden, zeigt dieses Gebilde eine Trägheit. Hinzu kommt, daß die Ausrichtung eines Dipols vor allem in kondensierten Stoffen durch seine Wechselwirkung mit der atomaren oder molekularen Umgebung behindert wird. Diese Eigenschaften führen dazu, daß ein permanenter Dipol sich zwar sehr schnell aber nicht schneller als mit $\approx 10^{-11}$ s einstellen kann. Die damit festliegende Einstellzeit bedeutet, daß permanente Dipole etwa jeder Wechselfrequenz bis 1 GHz ohne Verzögerung folgen können, daß sie etwa im Frequenzbereich 10 bis 1000 GHz in der Frequenzabhängigkeit von ε' und ε'' deutlich wirksam werden, aber höherfrequenten Wechselfeldern praktisch nicht mehr zu folgen vermögen. Der Ausrichtung permanenter Dipole wirkt in jedem Fall die statistische Wärmebewegung entgegen.

Die *Ionenpolarisation* tritt in heteropolaren Festkörpern auf. Induzierte Dipolmomente entstehen hier dadurch, daß die positiven und negativen Ionen durch ein elektrisches Feld etwas aus ihren normalen Gitterpositionen herausgeschoben werden. Dann fallen für jedes beliebige Volumenelement, das mehr als ein Ionenpaar umfaßt, der Schwerpunkt von positiver Gesamtladung und negativer Gesamtladung nicht mehr zusammen. Derartige Dipole werden schon während ihrer Entstehung ausgerichtet. Die Einstellzeiten liegen hier um 10^{-13} s. Das entspricht elektromagnetischen Schwingungen im Infrarotgebiet.

Wird die Elektronenhülle eines Atoms oder Moleküls unter der Einwirkung eines elektrischen Felds so deformiert, daß auch für den einzelnen Gitterbaustein ein induziertes Dipol-

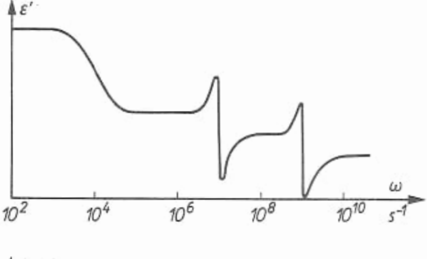
moment entsteht, wird von *Elektronenpolarisation* gesprochen. Für die leichten Elektronen betragen die Einstellzeiten nur etwa 10^{-15} s. Die mit der Elektronenpolarisation verbundenen Prozesse sind deshalb im sichtbaren Spektralgebiet und im UV-Gebiet nachweisbar. Noch höheren Wechselfrequenzen kann auch die Elektronenhülle nicht mehr folgen. Das heißt z. B., daß Röntgen- und Gammastrahlung keine Polarisation des Festkörpers bewirkt. Im positiven Sinne formuliert: mit Hilfe von Röntgenbeugungsexperimenten kann die ursprüngliche Elektronendichteverteilung in einem Festkörper nachgewiesen werden (↑ 6.4.).

Diese Erkenntnisse lassen sich in einer detaillierteren Formel zusammenfassen:

$$P = \sum_n N_{0n} p_n + \sum_n N_{In} \alpha_{In} E_{lok} + \sum_n N_{En} \alpha_{En} E_{lok} \quad (6)$$

(N_{0n} Zahl der permanenten Dipole der Sorte p_n, N_{In} Zahl der induzierten ionaren Dipole, N_{En} Zahl der induzierten elektronischen Dipole, α_{In} Polarisierbarkeit der Ionenkonfiguration n im lokalen Feld E_{lok}, α_{En} Polarisierbarkeit der Elektronenhülle der Bausteinsorte n). Abb. 8.6.2-1 faßt den durch die verschiedenen Polarisationsmechanismen bedingten Verlauf der Dielektrizitätskonstante und der dielektrischen Verluste zusammen.

8.6.2.3. Wagnersche Verluste. Auch in dem Frequenzbereich, in dem sich heute die Elektrotechnik und Elektronik vorwiegend betätigen, ist ein Dielektrikum nicht verlustfrei. Das läßt sich besonders einleuchtend an einem einfachen Beispiel zeigen, z. B. an einem Plattenkondensator, dessen Dielektrikum aus zwei Schichten besteht, die sich in ihrer Leitfähigkeit unterscheiden (Abb. 8.6.2-2). Für die Leitfähigkeit



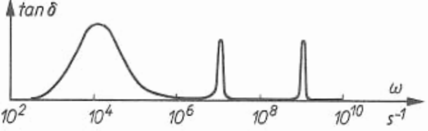


Abb. 8.6.2-1 Frequenzabhängigkeit der Dielektrizitätskonstante ε und der dielektrischen Verluste; tan δ Verlustwinkel

soll gelten $\sigma_1 > \sigma_2$. Dann folgt aufgrund der Kontinuitätsgleichung sofort für die Stromdichte $j = \sigma_1 E_1 = \sigma_2 E_2$ und $D_2 > D_1$. Auf der Grenzfläche zwischen beiden Dielektrika muß sich deshalb eine Ladungsmenge anhäufen; bei der hier gewählten Polarität, eine positive Ladungsschicht. Da in den Kondensator zusätzlich Ladung einfließt, ist $C_{eff} > C$ und auch

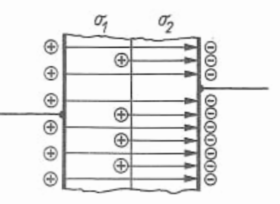


Abb. 8.6.2-2 Zweischichtdielektrikum

$\varepsilon_{eff} > \varepsilon$. Die Leitfähigkeit und die Ladungsträgerbeweglichkeit im Dielektrikum sind voraussetzungsgemäß sehr klein, und die notwendige Zusatzladung zur Einstellung des Gleichgewichts kann nicht ohne Verzögerung auf die Grenzfläche gelangen. Eine Abschätzung der dazu benötigten Zeit kann auf einfache Weise gewonnen werden. Man setzt einen Kondensator mit einer Schichtdicke von 10 µm voraus, nimmt einen typischen Meßwert für die Ladungsträgerbeweglichkeit $\mu = 10^{-6}$ cm²(Vs)⁻¹ an und arbeitet mit einer angelegten Spannung von 100 V. Dann gilt für die Transitzeit der Ladungsträger $\tau = d^2/U\mu = 10^{-2}$ s. Dieser Wert liegt im Frequenzbereich der Wagnerschen Verluste, der etwa von 1 Hz bis 10 kHz reicht. In realen Festkörperdielektrika sind Inhomogenitäten sowohl in der Leitfähigkeit als auch in der Dielektrizitätskonstanten nicht ganz zu vermeiden, was aber zum selben Resultat führt. Anders formuliert: die Technologie der Herstellung eines Dielektrikums muß eine möglichst einwandfreie Homogenität der Eigenschaften im gesamten Wirkvolumen sichern. Das muß z. B. auch an der Grenzfläche Si/SiO_2 für die Schaltkreisfertigung beobachtet werden, denn inzwischen ist klar geworden, daß an der Oxidgrenze eine sehr dünne Übergangszone SiO_x entsteht, in der sich x von kleineren Werten beginnend dem Wert 2 annähert. Durch Messung der Frequenzabhängigkeit des Verlustwinkels im angegebenen Bereich können Inhomogenitäten des Dielektrikums nachgewiesen werden.

8.6.2.4. Lokales Feld. Die Feldstärke, die mit E_{lok} bezeichnet wird und die den tatsächlichen Feldstärkewert im Inneren eines Dielektrikums angibt, ist nur in Ausnahmefällen gleich dem von außen angelegten Feld. Das liegt vor allem daran, daß jeder elektrische Dipol ein elektrisches Feld erzeugt. Ein Dipolmoment $\boldsymbol{p} = q\boldsymbol{s}$, dessen Ortsvektor $\boldsymbol{s}$ definitionsgemäß von der

negativen $-q$ zur positiven Ladung $+q$ weist, erzeugt um sich ein elektrisches Feld

$$(7) \quad \boldsymbol{E}(\boldsymbol{r}) = \frac{3(\boldsymbol{p}\boldsymbol{r})\boldsymbol{r} - r^2\boldsymbol{p}}{4\pi\varepsilon_0 r^5}.$$

Oft ist nur der Fall interessant, daß die Wellenlänge der mit dem Festkörper wechselwirkenden elektromagnetischen Welle viel größer als die Abmessung der Dipole bzw. irgendeiner anderen Ladungsverteilung ist. Dann genügt es, das lokale Feld in der räumlichen Mittelung $\boldsymbol{E} = \frac{1}{V}\int \boldsymbol{E}(\boldsymbol{r})\,\mathrm{d}V$ einzusetzen. Es gilt

$$(8) \quad \boldsymbol{E}_{\text{lok}} = \boldsymbol{E}_a + \boldsymbol{E}_1 + \boldsymbol{E}_2 + \boldsymbol{E}_3.$$

Hier ist $\boldsymbol{E}_a$ das von außen angelegte Feld, das man im Vakuum am gewählten Ort auch messen würde. $\boldsymbol{E}_1$ kommt durch die Polarisation $\boldsymbol{P}$ des Dielektrikums zustande und beträgt $\boldsymbol{E}_1 = -\gamma\boldsymbol{P}/\varepsilon_0 = -\boldsymbol{E}_a\gamma(\varepsilon - 1)$. Die Größe γ ist der Entelektrisierungsfaktor, der, wenn $\boldsymbol{P}$ homogen ist, nur von der Probenform abhängt. Eine homogene Polarisation ist nur für Dielektrika in Ellipsoidform erreichbar. Wichtige Fälle sind die Kugel mit $\gamma_x = \gamma_y = \gamma_z = 1/3$, der dünne Stab mit $\gamma_z = 0$, $\gamma_x = \gamma_y = 1/2$ und die dünne Platte oder Dünnschicht mit $\gamma_x = \gamma_y = 0$, $\gamma_z = 1$. $\boldsymbol{E}_1$ wird deswegen als Entelektrisierungsfeld bezeichnet. Das elektrische Feld innerhalb eines Dielektrikums ist immer kleiner als das von außen angelegte Feld, wenn $\varepsilon > \varepsilon_a$ gilt. Dabei ist ε_a die relative Dielektrizitätskonstante des das Dielektrikum umgebenden Außenraums. Hier wurde $\varepsilon_a = 1$ angenommen, was z. B. für Luft oder Vakuum gilt.

Schwerer einzusehen ist, daß zwei weitere Glieder $\boldsymbol{E}_2$ und $\boldsymbol{E}_3$ herangezogen werden müssen. Ein eleganter Weg zur Berechnung dieser Anteile stammt von LORENTZ (Abb. 8.6.2-3). Dabei geht es darum, die Fernordnung und Nahordnung (↑ 6.4.) des Dielektrikums gesondert in ihrer Wirkung auf einen bestimmten Punkt im Dielektrikum zu berücksichtigen. Man trennt um den Gitterpunkt herum ungefähr eine Kugel heraus, die 1000 Elementarzellen umfaßt, und beschreibt das außen liegende Gebiet mit der pauschalen Dielektrizitätskonstante ε. Die Oberfläche der gedachten Hohlkugel muß dann die Flächenladungsdichte $|\sigma| = \boldsymbol{P}\boldsymbol{r}/r$ besitzen, die das folgende Feld erzeugt:

$$(9) \quad |\boldsymbol{E}_2| = -\frac{1}{4\pi\varepsilon_0 r^2}\int |\sigma|\,\mathrm{d}A = \frac{P}{4\pi\varepsilon_0 r^2}\int_{-1}^{+1} r^2\, 2\pi \cos^2\alpha\,\mathrm{d}(\cos\alpha) = \frac{P}{3\varepsilon_0}.$$

$\boldsymbol{E}_3$ kann als Summe aller Dipolfelder der herausgetrennten Kugel berechnet werden. Für viele kubische Gitter ist $\boldsymbol{E}_3$ gleich Null. Da zur Berechnung des hier gewählten Beispiels $\varepsilon = \varepsilon_a$ vorgegeben ist, wird auch $\boldsymbol{E}_1 = 0$ und man erhält schließlich für ein kubisches Gitter

$$(10) \quad \boldsymbol{E}_{\text{lok}} = \boldsymbol{E}_a + \boldsymbol{P}/3\varepsilon_0.$$

8.6.2.5. Clausius-Mossotti-Gesetz. Für reine Elektronenpolarisation und kubisches Gitter folgt aus den Gleichungen 8.6.(6) und 8.6.(10) $\boldsymbol{P} = \sum_n N_{\text{En}}\alpha_{\text{En}}(\boldsymbol{E}_a + \boldsymbol{P}/3\varepsilon_0)$. Daraus ergibt sich $\varepsilon_0(\varepsilon - 1)\,E_a = \sum_n N_{\text{En}}\alpha_{\text{En}}E_a[1 + (\varepsilon - 1)\,\varepsilon_0/3\varepsilon_0]$ und das Gesetz von CLAUSIUS und MOSSOTTI

$$(\varepsilon - 1)/(\varepsilon + 2) = \sum_n N_{\text{En}}\alpha_{\text{En}}/3\varepsilon_0.$$



Abb. 8.6.2-3 Zur Berechnung des Lorentzfelds E_2; E_1 Elektrisierungsfeld

Für beliebige, kompliziertere Gitter müßte man den Wert $3\varepsilon_0$ rechts im Nenner durch einen anderen Ausdruck ersetzen. Diese Gleichung, die es gestattet, die relative Dielektrizitätskonstante eines Dielektrikums ohne permanente Dipole in einfacher Weise aus der Zahl der Bausteine N und deren Polarisierbarkeit α zu berechnen, kann auch zur Bestimmung des Temperaturkoeffizienten von ε benutzt werden. Man findet unter Berücksichtigung von $N_{\text{En}} = n_{\text{En}}/V$ den Ausdruck $(\mathrm{d}/\mathrm{d}T)\,(\varepsilon - 1)/(\varepsilon + 2) = \sum_n (\alpha_{\text{En}}/3\varepsilon_0)\,(\mathrm{d}/\mathrm{d}T)\,(n_{\text{En}}/V)$. Mit dem Volumenausdehnungskoeffizienten $\beta = \mathrm{d}V/V\mathrm{d}T$ gilt $[\mathrm{d}\varepsilon/(\varepsilon + 2)\,\mathrm{d}T] - [(\varepsilon - 1)/(\varepsilon + 2)^2] = -\beta\,[(\varepsilon - 1)\,/\,(\varepsilon + 2)]$. Schließlich ergibt sich für den Temperaturkoeffizienten

$$(11) \quad \mathrm{TK}_\varepsilon = \frac{1}{\varepsilon}\,\frac{\mathrm{d}\varepsilon}{\mathrm{d}T} = -\beta\,\frac{(\varepsilon - 1)\,(\varepsilon + 2)}{3\varepsilon}.$$

Da viele Dielektrika ε-Werte von 2 bis 3 aufweisen, ist $\mathrm{TK}_\varepsilon \approx -\beta$. Das ergibt größenordnungsmäßig eine Verringerung von ε um 1% bei einer Temperaturerhöhung um 100 K.

8.6.2.6. Relaxationserscheinungen. Polarisationsvorgänge, die auf der Einstellung permanenter Dipole beruhen, mit einer Veränderung des Lorentzfelds verbunden sind oder an inneren Grenzflächen ablaufen, benötigen aufgrund der dazu notwendigen Umordnung von Atomen oder Molekülen im Festkörper eine Relaxationszeit $\tau \approx 10^{-11}$ s. Nach dem Anlegen des elektrischen Felds zum Zeitpunkt $t = 0$ erhöht sich die Polarisation im Dielektri-

kum wie $P = P_0[1 - \exp(-t/\tau)]$, und nach Abschalten des Felds sinkt sie ab wie $P = P_0 \exp(-t/\tau)$. Einem elektrischen Wechselfeld kann die Polarisation nur verzögert folgen

$$P = (\varepsilon_\infty - 1)\,\varepsilon_0 E(t) + (\varepsilon_s - \infty)\,\varepsilon_0 \int_{-\infty}^{t} E(t\varepsilon') \exp[-(t - t')/\tau]\,\mathrm{d}t'/\tau.$$

Hierbei ist ε_∞ die relative Dielektrizitätskonstante im optischen Frequenzbereich. Sie charakterisiert die Polarisationsvorgänge, die keiner Relaxation unterliegen, weil sie dafür zu trägheitslos sind. Der Relaxation unterliegt deshalb nur der Anteil $(\varepsilon_s - \varepsilon_\infty)$, wobei ε_s die Dielektrizitätskonstante bei niedrigen Frequenzen darstellt. Der Integralterm beschreibt die verzögerte Einstellung der Dipole. In allen technisch wichtigen Fällen gilt für $E(t) = \sum E_n \exp \mathrm{i}\omega_n t$. Die Integration $\int_{-\infty}^{t} \exp \mathrm{i}\omega t' \exp t'/\tau\,\mathrm{d}t' = \exp \mathrm{i}\omega t \exp t/\tau/(\mathrm{i}\omega + 1/\tau)$ läßt sich leicht ausführen und man findet

$$(12)\quad P(t) = (\varepsilon_\infty - 1)\,\varepsilon_0 E_0 \sin \omega t + (\varepsilon_s - \varepsilon_\infty)\,\varepsilon_0 E_0 (\sin \omega t - \omega\tau \cos \omega t)/(1 + (\omega\tau)^2).$$

Die Anwendung trigonometrischer Formeln gestattet es, Gleichung 8.6.(12) in zwei Anteile zu zerlegen, in einen, der mit dem äußeren Feld in Phase ist, und in einen, der eine Phasendifferenz zum äußeren Feld aufweist:

$$(13)\quad P(t) = (\varepsilon_\infty - 1)\,\varepsilon_0 E_0 \sin \omega t + (\varepsilon_s - \varepsilon_\infty)\,\varepsilon_0 E_0 \sin(\omega t - \varphi)/(1 + (\omega\tau)^2)^{1/2}.$$

Die Phasendifferenz ist $\varphi = \arctan \omega\tau$. Setzt man voraus, daß die Leitfähigkeit des Dielektrikums verschwindend klein ist, dann genügt es, die Relaxationsstromdichte für die Berechnung des Verlustwinkels $\tan\delta$ heranzuziehen. Diese Stromdichte ist:

$$(14)\quad j_p(t) = \mathrm{d}P(t)/\mathrm{d}t = (\varepsilon_\infty - 1)\,\varepsilon_0 E_0\, \omega \cos\omega t + (\varepsilon_s - \varepsilon_\infty)\,\varepsilon_0 E_0 (\cos \omega t + \omega\tau \sin \omega t)/(1 + (\omega\tau)^2).$$

Der Verlustwinkel folgt dann als

$$(15)\quad \tan\delta = \frac{(\varepsilon_s - \varepsilon_\infty)\,\omega\tau}{(\varepsilon_s - \varepsilon_\infty) + \varepsilon_\infty(1 + (\omega\tau)^2)} = \frac{\varepsilon''}{\varepsilon'}.$$

Um dieses Ergebnis zu erhalten, wurde der Quotient aus dem Stromdichteanteil j_p gebildet, der mit der angelegten Feldstärke in Phase ist und dem Anteil, der eine Phasendifferenz von $\pi/2$ besitzt. Im Nenner wurde neben den mit einem cos-Faktor behafteten Gliedern in j_p der Anteil $j_E = \mathrm{d}E\varepsilon_0/\mathrm{d}t$, der Verschiebungsstrom, eingesetzt. Zur Definition von $\tan\delta$ sollte die Darstellung Abb. 8.6.2-4 herangezogen werden. Die Gleichungen 8.6.(13) und 8.6.(15) zeigen, daß Relaxationserscheinungen im Frequenzgebiet $\omega \approx 1/\tau$ auftreten. In diesem Bereich ändert sich die relative Dielektrizitätskonstante kontinuierlich von ihrem stationären Wert bis zum Wert im optischen Bereich. Der Maximalwert von $\tan\delta$ wird hier bei $\omega\tau = (\varepsilon_s/\varepsilon_\infty)^{1/2}$ erreicht. Eine Relaxationserscheinung wird in Abb. 8.6.2-5 dargestellt.

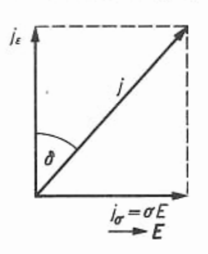


Abb. 8.6.2-4 Zur Definition des Verlustwinkels $\tan\delta$

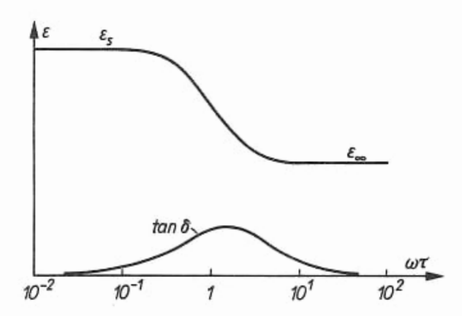


Abb. 8.6.2-5 Relaxationserscheinung in einem Dielektrikum

8.6.2.7. Resonanzerscheinungen. Wie bereits festgestellt, entstehen induzierte Dipole im optischen Frequenzbereich des elektromagnetischen Spektrums. Sowohl bei der Verschiebung von Ionen aus ihrer Ruhelage als auch bei der Deformation einer Elektronenhülle wirken der Kraft des elektrischen Wechselfelds Rückstellkräfte entgegen. Anhand eines einfachen Modells kann eine mathematische Formulierung des Vorgangs der Polarisation vorgenommen werden.

Ein elektrisches Feld $E_x = E_{x_0} \exp(\mathrm{i}\omega t)$ soll auf einen geladenen Baustein einwirken, der wie ein harmonischer Oszillator mit der Kraft $F = -c(x - x_0)$ an seine Ruhelage x_0 gebunden ist. Die bei der Verschiebung des Bausteins auftretenden Wechselwirkungen mit der Umgebung werden durch eine geschwindigkeitsproportionale Reibungskraft $F_r = -a\,\mathrm{d}x/\mathrm{d}t$ berücksichtigt. Unter diesen Voraussetzungen findet man die Bewegungsgleichung nach DRUDE

und LORENTZ für einen geladenen Baustein in einem lokalen elektrischen Wechselfeld:

(16) $m\,\mathrm{d}^2x/\mathrm{d}t^2 + a\,\mathrm{d}x/\mathrm{d}t + c(x - x_0)$
$= qE_{x,\mathrm{lok}}$.

Diese Gleichung läßt sich leicht in eine Beziehung umformen, die als Variable die Polarisation P enthält. Anknüpfend an Gleichung 8.6.(5) kann man schreiben: $P = N\lambda q(x - x_0)$. Mit Hilfe des konstanten Faktors λ sollen die Nahordnung und Fernordnung des konkreten Gitters in ihrer Wirkung auf die sich wirklich einstellende Größe des induzierten Dipolmoments $\lambda q(x - x_0)$ beachtet werden. N ist die Zahl der sich einstellenden Bausteine pro Volumeneinheit. Also ergibt sich $m\,\mathrm{d}^2P/\mathrm{d}t^2 + a\,\mathrm{d}P/\mathrm{d}t + cP = \lambda Nq^2E_{x,\mathrm{lok}} = \lambda Nq^2(E_{0x}\exp(\mathrm{i}\omega t) + P/3\varepsilon_0)$. Die Lösung der Gleichung lautet

(17) $P = E_{0x}\exp(\mathrm{i}\omega t)\,B/(\omega_0^2 - \omega^2 + \mathrm{i}a\omega/m)$.

In dieser Abhängigkeit der Polarisation P von der Frequenz ω drückt sich deutlich der Resonanzcharakter des induzierten Polarisationsvorgangs aus. In 8.6.(17) ist B eine Konstante und $\omega_0^2 = c/m - \lambda Nq^2/3\varepsilon_0 m$. Die Frequenz ω_0 ist die Eigenfrequenz des geladenen Bausteins im Gitter eines bestimmten Festkörpers. Sie liegt für die Ionenpolarisation im Infrarotgebiet und für die Elektronenpolarisation im UV-Bereich. Die Polarisation kann dem äußeren Feld nur verzögert folgen. Abb. 8.6.2-6 zeigt eine Resonanzerscheinung in ihren charakteristischen Abhängigkeiten der komplexen Dielektrizitätskonstanten ε und des Verlustwinkels $\tan\delta$ von der Frequenz.

8.6.3. Leitungsverhalten von Isolierstoffen

8.6.3.1. Ionenleitung. Ionenleitung kann in Festkörpern in der Form der Eigenleitung auftreten, wenn es sich z. B. um einen heteropolaren Festkörper handelt, in dem genügend Ionenbausteine für den Ladungstransport zur Verfügung stehen. Derartige normalerweise örtlich fixierte Bausteine können sich aufgrund von Platzwechselvorgängen durch das Gitter be-

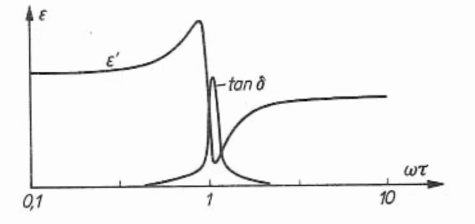


Abb. 8.6.2-6 Resonanzerscheinung in einem Dielektrikum

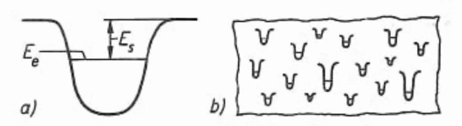


Abb. 8.6.3-1 Zur Hoppingleitung; a) Potentialbarriere, b) Verteilung von Potentialbarrieren verschiedener Tiefe in einem Volumenelement; E_s Höhe der Potentialbarriere

wegen. Dieser Vorgang wird bei hohen Temperaturen begünstigt. Viel weiter verbreitet ist der Ladungstransport in Isolierstoffen durch Verunreinigungen, die in Ionenform vorliegen. In beiden Fällen wird der Transport meßbarer Ladungsmengen erst durch die Defektstruktur des Gitters ermöglicht. So kann man davon ausgehen, daß sich Frenkel- oder Schottkydefekte unter der Wirkung eines elektrischen Felds auf der Grundlage des Platzwechselmechanismus durch ein heteropolares Gitter hindurchbewegen können.

8.6.3.2. Leitungsvorgänge bei niedrigen Feldstärken. Da in Isolierstoffen die Beweglichkeit der Ladungsträger meist sehr niedrig ist, sind offensichtlich Ladungstransportmechanismen vorherrschend, bei denen die Elektronen oder Defektelektronen während ihrer Lebensdauer im wesentlichen lokalisiert sind und sich nur in einem kleinen Bruchteil dieser Zeitdauer unter Feldeinfluß wirklich bewegen. Ein sehr verbreiteter Ladungsvorgang ist deshalb die *Hoppingleitung*. Durch eine intensive Elektron-Phonon-Wechselwirkung bleiben die Elektronen lange Zeit fest an einem Gitterplatz verankert. Die kleine Potentialmulde, in der sich der Ladungsträger aufhält, kann durch eine Potentialbarriere beschrieben werden, die verhindert, daß der Ladungsträger diesen Ort schnell verlassen kann. Eine geringe Wahrscheinlichkeit für das Entkommen aus diesem Ort ist durch die thermischen Stöße des Gitters gegeben. Charakteristisch für die Hoppingleitung ist eine *thermisch aktivierte Beweglichkeit* $\mu \sim \exp(-E_s/k_BT)$. E_s ist die Höhe der Potentialbarriere, gerechnet von der Elektronenenergie bis zur Oberkante der Potentialbarriere (Abb. 8.6.3-1). Oft wird allerdings der viel allgemeinere Fall beobachtet, daß in einem Isolierstoff eine Fülle verschieden stark lokalisierter Plätze im Volumen des Festkörpers verstreut liegen. Dann ist E_s das Ergebnis einer Überlagerung verschiedener Potentialtiefen und temperaturabhängig. Bei genaueren Überlegungen ist zu berücksichtigen, daß der „hüpfende" Ladungsträger auch einen geeigneten „Landeplatz" benötigt.

Betrachtet man den Vorgang der Hoppingleitung in einem heteropolaren Gitter näher, stellt sich heraus, daß der lokalisierte Ladungsträger seine Gitterumgebung polarisiert. Es läßt sich

zeigen, daß der polarisierte Nahordnungsbereich zusammen mit dem Elektron durch das Gitter hindurch bewegt wird. Das Quasiteilchen aus Elektron und Polarisationszustand bezeichnet man als *Polaron*. Polaronen haben infolge ihrer Wechselwirkung mit dem Gitter Massen von etwa 1,5 bis 6 Elektronenmassen.

Von MOTT stammt ein Modell, das es ermöglicht, viele Besonderheiten des Ladungstransports in Isolierstoffen zu erfassen. Es beruht darauf, daß das Bändermodell, wie Abb. 8.6.3-2

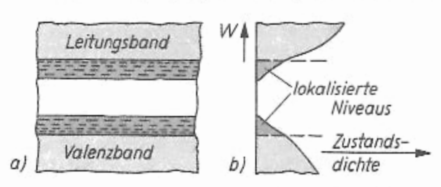


Abb. 8.6.3-2 Mottmodell; a) Bänder mit vorgelagerten Elektronenniveaus, b) Zustandsdichte

zeigt, durch lokalisierte Niveaus in der verbotenen Zone erweitert wird. Nach wie vor können sich Ladungsträger entsprechender Energie in den Bändern aufhalten und dort mit relativ hoher Beweglichkeit bewegen. Dagegen sind Ladungsträger, die sich innerhalb der verbotenen Zone aufhalten, zunehmend mit ihrem energetischen Abstand von der Bandkante unbeweglicher. Das Modell geht davon aus, daß die Dichte der den Bandkanten vorgelagerten Niveaus mit dem Abstand von der Bandkante abnimmt. Gleichermaßen steigt die Lokalisation der Ladungsträger und ihre Beweglichkeit nimmt stark ab. Bezogen auf die Gesamtzahl der zur Verfügung stehenden Ladungsträger bildet sich eine effektive Ladungsträgerbeweglichkeit heraus.

Die energetisch bedingten Plätze, in denen sich Ladungsträger längere Zeit aufhalten können, sind die *Haftstellen*. Als *flache Haftstellen* bezeichnet man solche, die in unmittelbarer Nachbarschaft der Bandkanten liegen und deshalb thermisch leicht in das Band entleert werden können. *Tiefe Haftstellen* dagegen halten Ladungsträger für lange Zeiten fest. Für sie ist die Entleerung ein seltener Vorgang. Bezeichnet man mit dem Index f die flachen Haftstellen und mit t die tiefen, dann gelten für die Abstände der Haftniveaus von der Bandkante ΔE und für die Anteile der Ladungsträger in den verschiedenen Zuständen, wobei n die Konzentration im Leitungsband sein soll, die typischen Ungleichungen $k_B T \approx \Delta E_f \ll \Delta E_t$ und $n \ll n_f \ll n_t$. Das *Mottmodell* kann auch dadurch beschrieben werden, daß man es als ein Bändermodell mit unscharfen Bandkanten auffaßt.

8.6.3.3. Leitungsvorgänge bei hohen elektrischen Feldstärken. Während in der Praxis bei Isolierproblemen die Wirkungen kleiner elektrischer Felder kaum ins Gewicht fallen, ist der Einfluß starker elektrischer Felder sehr beachtlich. Zunächst sei aber einiges zu den Leitungsprozessen gesagt, die zwar höhere Ströme zulassen, aber noch nicht zu irreversiblen Änderungen in der Struktur der Isolierstoffe führen. Unter der Wirkung hoher elektrischer Felder werden an den Elektroden oder im Inneren des Isolators bestehende Potentialbarrieren erniedrigt. Das geht bei entsprechenden Feldstärken soweit, daß die hemmende Wirkung der Barriere fast völlig verschwindet. Die starke Bereitstellung von Ladungsträgern an einer Elektrode durch Feldeinfluß bezeichnet man als *Feldemission*. Auch thermisch können durch Emission hohe Ladungsträgerkonzentrationen an einer Metallelektrode bereitgestellt werden (Schottkyemission, ↑ 8.3.6.2.9.). Grenzt das Metall an einen Isolator, so erniedrigt sich die Austrittsarbeit infolge der Wirkung der Bildladung an der Grenzfläche von E_A zu $E_A - (e^3 E/4\pi\varepsilon\varepsilon_0)^{1/2}$.

Man kann generell feststellen, daß eigentlich immer genügend Ladungsträger für die Injektion in den Isolierstoff an den Elektroden zur Verfügung stehen. Wichtig ist es deshalb, die Mechanismen näher zu beleuchten, die den Stromtransport durch einen Isolator hindurch oder auf seiner Oberfläche gesetzmäßig begrenzen. Ein derartiger Mechanismus ist der *Poole-Frenkel-Effekt*. Durch starke elektrische lokale Felder werden die Haftstellentiefen in einem Isolator stark verringert. Eine Haftstellentiefe ohne Feldeinfluß ΔE_H reduziert sich bei Vorhandensein von einem lokalen Feld $|E_{lok}|$ auf $\Delta E_H - (e^3 |E|_{lok}/\pi\varepsilon\varepsilon_0)^{1/2}$. Da in Haftstellen entsprechender Tiefe ausreichende Mengen von Ladungsträgern vorhanden sind, führt die Umwandlung tiefer Haftstellen in flache unter Feldeinfluß zu einem Anstieg der Stromdichte im Isolator.

Jede in Haftstellen eingefangene Ladungsträgerkonzentration, die innerhalb des Isolators räumlich verteilt ist, bildet eine *Raumladung*. Es muß allerdings darauf hingewiesen werden, daß auch ohne Haftstellen in einem Isolator eine Raumladung entsteht. Für den *haftstellenfreien Isolator* findet man für den Stromfluß das *Childsche Gesetz*

$$(18) \quad j = (9/8)\, \varepsilon\varepsilon_0 \mu U^2/d^3 .$$

U ist die an einer Sandwichanordnung anliegende Spannung und d die Dicke des Isolators. Zwei wichtige Fälle des Haftstelleneinflusses sollen hier besonders hervorgehoben werden. Erstens kann man voraussetzen, daß irgendein Mechanismus zu einem Ohmschen Stromfluß Anlaß gibt und ein diskretes Haftstellenniveau der Tiefe ΔE_H im Isolator vorliegt. Dann findet man ein *modifiziertes Childsches Gesetz*:

(19) $j = en\mu U/d + (9/8)\,\varepsilon\varepsilon_0 \Theta U^2/d^3$.

Θ ist hier das Verhältnis der Ladungsträger im Leitungsband zur Gesamtzahl aller injizierten Ladungsträger. Dieser Faktor errechnet sich aus der Zustandsdichte im Leitungsband N_c und der Konzentration der Haftstellen N_H nach $\Theta = (N_c/N_H) \exp(-\Delta E_H/k_B T)$. Dabei muß vorausgesetzt werden, daß es sich um tiefe Haftstellen handelt.

Zweitens kommt in der Praxis recht häufig der Spezialfall vor, daß die Haftstellen in der verbotenen Zone energetisch exponentiell verteilt sind. Eine derartige Verteilung kann mit $h(E) = (H/k_B T_c) \exp[(E - E_c)/k_B T_c]$ beschrieben werden. Hierin bedeuten T_c ein Parameter, der die Steilheit der exponentiellen Verteilung charakterisiert, E_c die Unterkante des Leitungsbands oder die Lage der Beweglichkeitskante und H die Gesamtkonzentration aller Haftstellen innerhalb der verbotenen Zone. Die Abhängigkeit der Stromdichte von der Spannung U ist durch einen überproportionalen Anstieg charakterisiert, wobei $l = T_c/T$ eine Zahl bezeichnet, die Werte von 1,5 bis 12 haben kann. $f(l)$ ist eine Zahl mit der Größenordnung Eins. Die Stromdichte ist hier

(20) $j = e\mu N_c (\varepsilon\varepsilon_0/eH)^l f(l)\, U^{l+1}/d^{2l+1}$.

Die sich ergebenden Strom-Spannungs-Charakteristiken zeigt Abb. 8.6.3-3.

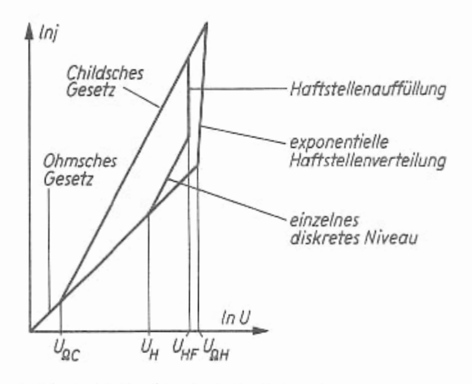


Abb. 8.6.3-3 Raumladungsbeschränkte Ströme, Strom-Spannungs-Charakteristiken; j Stromdichte

8.6.3.4. Durchschlagsmechanismen. Es ist von großem praktischem Wert, die Zerstörung von Isolierkörpern und Isolierschichten zu vermeiden, wie sie durch das Anlegen zu hoher Spannungen auftreten kann. Die Frage nach dem sicheren Einsatz von Isolierstoffen ist nicht leicht zu beantworten, weil viele Parameter die höchste zulässige Spannung beeinflussen. Hier müssen vor allem der Einfluß der Temperatur, der Luftfeuchtigkeit, des Elektrodenmaterials aber auch der von Strahlungen, von Gasen und Dämpfen und von Verschmutzungen hervorgehoben werden.

Der Durchschlagsvorgang selbst kann weit weniger gut formelmäßig erfaßt werden, weil er keinesfalls ein homogen ablaufender Vorgang ist. Vielmehr bilden sich meist kleine Durchschlagskanäle aus, deren Zahl und Größe mit wachsender Feldstärke zunimmt. Für jeden einzelnen Durchschlagskanal kann man grundsätzlich zwei Stadien unterscheiden:

Zunächst wächst die Stromdichte, ausgehend von einem der soeben beschriebenen Mechanismen mit überproportionalem Anstieg, sprunghaft an. Dadurch wird der Isolator mechanisch oder thermisch am Ort des Durchschlags zerstört. Die Zerstörung kann mit einem Ausheileffekt bezüglich des Isolationsvermögens der gesamten Anordnung verbunden sein. Es kann also örtlich ein Durchschlagskanal entstehen, der nur kurzzeitig zu einer Stromspitze führt. Oft liegt das einfach daran, daß durch Verdampfungen oder explosionsartiges Wegschleudern die äußere Stromzuführung an der Durchschlagsstelle verschwindet.

Durchschläge werden bezüglich der Durchschlagsursache in drei Gruppen eingeteilt. Beim *thermischen Durchschlag* entsteht durch eine örtliche Zunahme der Stromdichte eine spezifisch höhere Freisetzung von Joulescher Wärme, was sukzessive zu einer noch höheren Stromdichte und einem örtlichen Wärmestau führt, der schließlich die örtliche Zerstörung des Isolators bewirkt. Dieser Vorgang kommt zum Stehen, wenn schließlich auch das benachbarte Elektrodengebiet mit zerstört wird. Der *elektrische Durschlag* erfolgt dadurch, daß auf rein elektronischem Wege plötzlich lawinenartig eine große Anzahl von Ladungsträgern bereitgestellt wird. Unter dem dritten Typ, dem *elektromechanischen Durchschlag*, versteht man die Zerstörung des Isolatormaterials durch die starken elektrostatischen Anziehungskräfte zwischen geladenen Metallelektroden. Diese letztgenannte Art hat vor allem dort Bedeutung, wo es sich um relativ weiche Materialien wie Plastwerkstoffe handelt.

8.6.4. Ferroelektrika und Antiferroelektrika

8.6.4.1. Typische Eigenschaften der Ferroelektrika. Die Charakterisierung eines typischen Dielektrikums (↑ 8.6.1.2.) kann hier voll übernommen werden, muß aber ergänzt werden, weil in Ferroelektrika eine spontane Polarisation auftritt. Folgendes kann besonders hervorgehoben werden:

Die spontane Polarisation tritt nur in der Richtung ganz bestimmter kristallographischer Achsen auf. Unter der Wirkung eines ausrei-

chend starken elektrischen Felds ist die Polarisation in ihrer Richtung umkehrbar. Wird die Polarisationsrichtung einmal von einer bestimmten Orientierung aus beginnend umgepolt und dann wieder in die ursprüngliche Richtung gebracht, so durchläuft die Abhängigkeit der Polarisation von der Feldstärke eine Hysteresekurve. Oberhalb einer charakteristischen Umwandlungstemperatur gehen die meisten Ferroelektrika in den parelektrischen Zustand über. In unmittelbarer Nachbarschaft der Umwandlungstemperatur besitzt die relative Dielektrizitätskonstante eines Ferroelektrikums ein ausgeprägtes Maximum, das bis in die Größenordnung 10^5 reichen kann.

Ferroelektrika zeigen gleichzeitig Erscheinungen der Pyro- und Piezoelektrizität. In der Nähe der Umwandlungstemperatur ordnet ein ferroelektrischer Festkörper sein Gitter in der Nahordnung aber zumeist auch in der Fernordnung um, wobei die bei höherer Temperatur stabile Phase die größere Symmetrie aufweist. Ferroelektrische Einkristalle können gleichzeitig eine völlig einheitliche Polarisationsrichtung haben (Eindomänenkristall) oder als Mehrdomänenkristall auftreten. In der Umgebung der Umwandlungstemperatur werden an ferroelektrischen Festkörpern anomale Temperaturabhängigkeiten der piezoelektrischen und elastischen Koeffizienten, der spezifischen Wärmekapazität, der Leitfähigkeit sowie einiger optischer Eigenschaften gefunden.

8.6.4.2. Ferroelektrische Stoffe. Bisher gibt es keinen völlig befriedigenden Vorschlag, die Ferroelektrika nach bestimmten Prinzipien in Gruppen einzuteilen. Hier erfolgt die Einteilung nach dem Vorhandensein oder Fehlen eines Symmetriezentrums des Gitters. Der Typ der *Ferroelektrika mit Symmetriezentrum* umfaßt Bariumtitanat ($BaTiO_3$), Bleititanat ($PbTiO_3$), Triglycinsulfat $[(NH_2CH_2COOH)_3 \cdot H_2SO_4]$, Triglycinselenat $((NH_2CH_2COOH)_3 \cdot H_2SeO_4)$, Lithiumniobat ($LiNbO_3$), Cadmiumniobat ($Cd_2Nb_2O_7$), Bleiniobat ($PbNb_2O_6$), Colemanit ($Ca_2B_4O_{11} \cdot 5\,H_2O$) und andere. Zum Typ *ohne Symmetriezentrum* gehören unter anderen Seignettesalz ($NaKC_4H_4O_6 \cdot 4\,H_2O$), Kaliumdihydrogenphosphat (KH_2PO_4, KDP), Rubidiumdihydrogenphosphat (RbH_2PO_4), Caesiumdihydrogenarsenat (CsH_2AsO_4). Weitere wichtige Ferroelektrika, die sich hier nicht einordnen lassen, sind Guanidinaluminiumsulfathexahydrat $\{[C(NH_2)_3]\,[Al(H_{2O})_6]\,[(SO_4)_2]\}$, bestimmte Alaune, Sulfate und Selenate.

8.6.4.3. Wichtige Eigenschaften der Ferroelektrika. GINZBURG wandte die allgemeine Londonsche Theorie der Phasenumwandlung auf die Ferroelektrika an und erhielt für die freie Enthalpie

$$(21)\quad G = G_0 + a(T - T_c)\,P^2 + bP^4 + cP^6,$$

also einen viergliedrigen Ausdruck in geraden Potenzen der *Polarisation P*. In G_0 sind alle hier nicht weiter interessierenden Enthalpieanteile des Ferroelektrikums zusammengefaßt. Die Größen a, b und c sind Konstanten, die für einen bestimmten Stoff charakteristisch sind. Gleichung 8.6.(21) läßt sich für Ferroelektrika verwenden, die nur eine einzige polare Achse besitzen. Mit Hilfe dieser grundlegenden Beziehung lassen sich einige wichtige Zusammenhänge errechnen. Man erhält z. B. für die Abhängigkeit der Polarisation von der Feldstärke $\partial G/\partial P = |E| = 2a(T - T_c)\,P + 4bP^3 + 6cP^5$. Hier wurde davon ausgegangen, daß die freie Enthalpie als eine Energiegröße für ein Ferroelektrikum den Anteil $G' = EP$ enthält und dieser die drei Glieder mit P in 8.6.(21) umfaßt. Aus der abgeleiteten Beziehung $E(P)$ kann zumindest graphisch immer der Verlauf $P(E)$ ermittelt werden. Führt man das mit dem Experiment angepaßten Konstanten a, b und c aus, ergibt sich prinzipiell der Verlauf einer Hysteresekurve (↑ Abb. 8.6.4-4).

Mit Hilfe dieser phänomenologischen Theorie läßt sich auch der Verlauf der Dielektrizitätskonstanten in der Umgebung des ferroelektrischen Umwandlungspunkts auffinden. Ausgangspunkt ist dabei Gleichung 8.6.(4) $P = (\varepsilon - 1)\,\varepsilon_0 E$. Deshalb muß gelten
$1/(\varepsilon - 1) = 1/\varkappa_{el} = \varepsilon_0 \partial E/\partial P = \varepsilon_0\,\partial^2 G/\partial P^2$,
$1/(\varepsilon - 1) = \varepsilon_0(2a(T - T_c) + 12bP^2 + 30cP^4)$.
Daraus folgt für die nichtpolare Phase, oberhalb der Umwandlungstemperatur T_c, für die die spontane Polarisation gleich Null ist,

$$(22)\quad \varepsilon - 1 = \varkappa_{el} = C/(T - T_c).$$

Das ist das bekannte *Curie-Weiß-Gesetz*, das den Verlauf der Dielektrizitätskonstanten oberhalb der Umwandlungstemperatur richtig beschreibt. Eine genauere Rechnung zeigt, daß ε nicht unendlich wird. C ist in dem Modellfall $C = 1/2a\varepsilon_0$ und wird als *Curiekonstante* bezeichnet. Bei Phasenumwandlungen (ferroelektrisch – parelektrisch) vollzieht $\varepsilon - 1$ am Umwandlungspunkt einen Sprung von $2c/b^2\varepsilon_0$ (parelektrische Seite) nach $c/2b^2\varepsilon_0$, d. h., der Wert von $\varepsilon - 1$ wird bei Überschreiten der Umwandlungstemperatur in Richtung des ferroelektrischen Zustands um 1/4 reduziert.

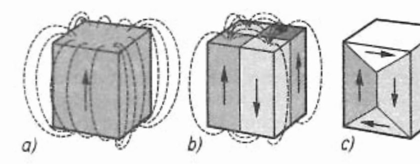


Abb. 8.6.4-1 Äußere Feldverteilung bei verschiedenen Domänenanordnungen

8.6.4.4. Ferroelektrische Domänen. Eine *Domäne* in einem Festkörper ist ein bestimmtes räumliches Gebiet, in dem eine bestimmte Zustandsgröße, meist vektoriellen Charakters, überall den gleichen Wert hat. Eine ferroelektrische Domäne ist ein Raumgebiet eines Dielektrikums, das eine überall gleichgroße spontane Polarisation gleicher kristallographischer Richtung besitzt.

Normalerweise besteht ein ferroelektrischer Kristall aus mehreren Domänen, weil dieser Zustand energetisch günstiger ist (Abb. 8.6.4-1). Zunächst läßt Gleichung 8.6.(21) die Aussage zu, daß eine spontane Polarisation $+P_s$ einer entgegengesetzt gerichteten Polarisation $-P_s$ energetisch völlig gleichwertig ist. Diese Eigenschaft des ferroelektrischen Zustands läßt also eine beliebige Einteilung eines einachsigen Ferroelektrikums in antiparallel orientierte Domänen zu. Neben dem durch Gleichung 8.6.(21) gegebenen Energieanteil, den man als eine Energie charakterisieren kann, die im Volumen einer Domäne homogen enthalten ist, sind noch weitere Energieanteile zu berücksichtigen, wenn die Gleichgewichtseinstellung eines Ferroelektrikums betrachtet werden soll. Wichtig sind für eine solche Betrachtung, bei der es letztlich um die Suche nach dem Zustand mit minimaler freier Enthalpie geht, die Energieanteile. Elektrostatische Energie im Außenraum um den ferroelektrischen Festkörper kann praktisch gleich Null werden, wenn die Domäneneinteilung so erfolgt, daß keine Feldlinien aus dem Festkörper austreten können (↑ Abb. 8.6.4-1 c).

Wandenergien der Domänenwände sind an den Grenzflächen lokalisiert, wo zwei verschieden orientierte Polarisationsrichtungen im Festkörper zusammentreffen. Eine *Domänenwand* in einem Ferroelektrikum hat eine Dicke von etwa 2 bis 10 nm (Abb. 8.6.4-2). Innerhalb dieser Wanddicke ändert sich die Größe der spontanen Polarisation z. B. in einem einachsigen Ferroelektrikum von $+P_s$ über ± 0 zu $-P_s$. Den Ursprung für eine erhöhte Konzentration von Energie im Wandbereich bilden mechanische Deformationen des Kristalls in dieser Region.

Die *Domänenstruktur* eines Realkristalls bildet sich zunächst zufällig bei seiner Entstehung heraus. Meist stellt sich die spontane Polarisation in einem größeren Kristall so ein, daß ein Multidomänenkristall entsteht. Wird dieser Kristall allerdings mit Elektroden versehen, bildet sich ein Eindomänenkristall. Das liegt daran, daß durch die Metallelektroden das Außenfeld abgebaut wird und es dann günstiger ist, auch die in den Domänen gespeicherte Energie abzubauen.

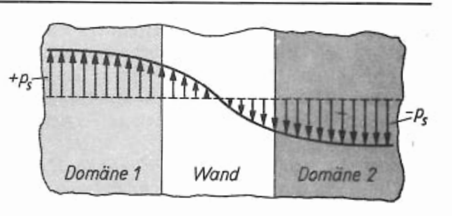


Abb. 8.6.4-2 Domänenwand eines Ferroelektrikums

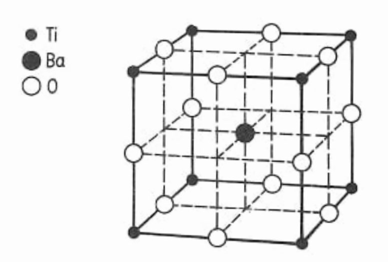


Abb. 8.6.4-3 Struktur des Bariumtitanats

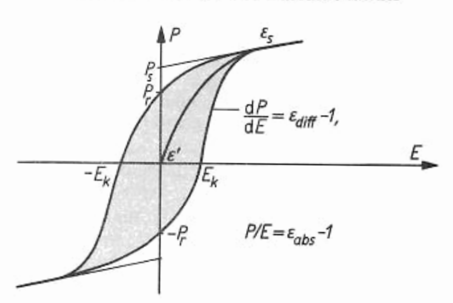


Abb. 8.6.4-4 Ferroelektrische Hysteresekurve

8.6.4.5. Bariumtitanat. Obwohl inzwischen mehr als 100 Ferroelektrika bekannt geworden sind, die in Kondensatoren, für Piezobauelemente, als Kaltleiter, als Varikonden, für elektrooptische Wandler, als pyroelektrische Strahlungsempfänger und in der nichtlinearen Optik Verwendung finden, sollen hier nur die Eigenschaften von Bariumtitanat abgehandelt werden.

Bariumtitanat wurde als Ferroelektrikum 1945 unabhängig voneinander von zwei Forschergruppen in der Sowjetunion (Vul und Goldman) und in den USA (Hippel, Breckenridge, Chesley und Tisza) erkannt. Mit dieser Substanz konnten erstmals die schon beschriebenen Zusammenhänge erklärt werden, da sie in gut überschaubarer Weise das Wesen eines Ferroelektrikums zeigt. Bis dahin war in der Untersuchung und in der Anwendung das 1655 von Seignette synthetisierte *Seignettesalz* führend gewesen. 1880 wurde die Piezoelektrizität und 1921 die Ferroelektrizität des Seignettesalzes entdeckt.

Bariumtitanat besitzt Perowskitstruktur (Abb. 8.6.4-3). Es wird als Bariumtitanatkeramik oder

in Form von Einkristallen verwendet. Einkristalle werden nach der Stockbargermethode gewonnen (↑ 6.4.2.1.). Durch geeignete Kristallzüchtungsverfahren erhält man Eindomänenkristalle, die eine nahezu rechteckige Hysteresekurve aufweisen.

Abb. 8.6.4-4 zeigt eine *ferroelektrische Hysteresekurve*, wie sie für Bariumtitanat typisch ist (magnetische Hysterese ↑ 8.7.3.). Eine derartige Hysteresekurve ist durch die Parameter *remanente Polarisation* P_r, *Sättigungspolarisation* und *Koerzitivfeldstärke* $|E_k|$ charakterisiert. Soll eine gegebene Polarisationsrichtung in Bariumtitanat umgekehrt werden, so ist mindestens die Feldstärke $|E_k|$ dazu erforderlich. Weiterhin läßt sich ein Ferroelektrikum durch bestimmte Werte der relativen Dielektrizitätskonstanten näher bestimmen, die man der Abhängigkeit $P(|E|)$ entnehmen kann. Im Beispiel sind das die Dielektrizitätskonstante bei kleinem Feld ε', die *differentielle Dielektrizitätskonstante* $\varepsilon_{\text{diff}}$, die *absolute Dielektrizitätskonstante* ε_{abs} und der Sättigungswert der Dielektrizitätskonstanten. Der Verlauf einer Hysteresekurve läßt sich damit erklären, daß für einen vollen Umlauf Energie verbraucht wird. Diese Energie wird benötigt, um die atomaren Bausteine geringfügig gegenüber ihrer Ruhelage zu verschieben, um ganze Domänen umzupolen oder um Domänenwände zu verschieben. Diese Vorgänge sind nicht ohne Reibungsverluste möglich, vor allem deshalb nicht, weil auch ein ferroelektrischer Kristall eine Defektstruktur besitzt. Insbesondere können sich Domänenwände in Gitterdefekten, wie Punktdefekten und Versetzungen, verhaken und sind dann nur unter Anwendung sehr hoher Feldstärken noch weiter zu verschieben. Der Energiebedarf für einen Umlauf, der letztendlich in Wärme übergeht, kann der Hysteresekurve als $\iint \boldsymbol{P} \mathrm{d}\boldsymbol{E}\, \mathrm{d}V$ entnommen werden. Auch bei sehr hohen Feldstärken ist noch ein Anwachsen der Polarisation zu verzeichnen. Das rührt von der zusätzlichen Entstehung induzierter Dipole her. Mit Hilfe von Neutronenbeugungsexperimenten konnte gezeigt werden, daß der Polarisationsvorgang in Bariumtitanat durch die räumliche Verschiebung von Gitterbausteinen in neue Gleichgewichtslagen erfolgt. Geht man von der nichtpolaren kubischen Phase aus, dann verschieben sich beim Übergang in die polare Phase die Titaniumatome um $6 \cdot 10^{-12}$ m in die eine und die auf verschiedenen kristallographischen Positionen sitzenden Sauerstoffatome um $5 \cdot 10^{-12}$ m bzw. $9 \cdot 10^{-12}$ m in die entgegengesetzte Richtung. Dadurch bildet sich in jeder Elementarzelle ein elektrisches Dipolmoment.

Bariumtitanat ändert reversibel bei Abkühlung dreimal seine Kristallstruktur. Es geht bei 120 °C vom kubischen System in das tetragonale System über, bei −5 °C erfolgt eine Umwandlung in die orthorhombische Phase und schließlich findet ein dritter Modifikationswechsel bei −90 °C statt. Auch diese Phasenumwandlungen sind nicht hysteresefrei. Unterhalb 120 °C bleibt $BaTiO_3$ ferroelektrisch, aber bei jedem Modifikationswechsel treten Anomalien im Verlauf bestimmter Eigenschaften auf. Das ist z. B. besonders ausgeprägt für die relative Dielektrizitätskonstante. Einige Angaben über Ferroelektrika sind in Tabelle 8.6.4-5 zusammengestellt.

Tab. 8.6.4-5 Eigenschaften der Ferroelektrika; T_C Curietemperatur in °C, ε Dielektrizitätskonstante, C Curiekonstante in °C

Ferroelektrikum	T_C	ε	C
$BaTiO_3$	120	10^4	$1{,}7 \cdot 10^4$
TGS	49	$6 \cdot 10^3$	$3{,}3 \cdot 10^3$
$Cd_2Nb_2O_7$	−90	$5 \cdot 10^3$	$4{,}6 \cdot 10^4$
KDP	−150	10^5	–

8.6.4.6. Sichtbarmachen von Domänenstrukturen. Für die Charakterisierung eines Ferroelektrikums ist es wichtig, seine Domänenstruktur zu kennen. Diese hängt von äußeren Einflüssen und elektrischen Feldern ab.

Mit Hilfe der *Ätztechnik* (Tafeln 38/39) kann eine momentane Domänenkonfiguration dauerhaft fixiert werden. Durch Anätzen der Kristalle kann die Verteilung von positiven und negativen Endflächen von Domänen auf der Kristalloberfläche sichtbar gemacht werden, weil sich die in entgegengesetzter Richtung polarisierten Kristallgebiete in geeigneten Lösungsmitteln verschieden schnell auflösen. Zum Beispiel lösen sich die positiven Endflächen von C-Domänen von Bariumtitanat in Säuren schneller auf als die negativen.

Durch polarisationsmikroskopische Untersuchungen lassen sich verschieden orientierte Domänen feststellen, weil die spontane Polarisation optische Anisotropien hervorruft, die sichtbar gemacht werden können. Zwischen gekreuzten Polarisatoren lassen sich deshalb für durchsichtige Ferroelektrika Domänenstrukturen sichtbar machen. Da sich bestimmte Staubteilchen, wie Schwefelpulver oder pulverförmiges Pb_3O_4, aufgrund ihrer eigenen Ladung an negativen Domänenenden abscheiden, ist auch aus der Bedeckung oder Nichtbedeckung von Kristallflächen eine Rekonstruktion der momentanen Domänenstruktur möglich. Domänenwände in einem Ferroelektrikum sind aufgrund ihrer besonderen Struktur Orte hoher spontaner elektrischer Felder. An derartigen Stellen des Festkörpers kondensiert bevorzugt Wasserdampf. Dieser als *Taumethode*

bezeichnete Nachweis ergänzt in anschaulicher Weise die anderen Untersuchungsverfahren.

8.6.4.7. Antiferroelektrika. In Analogie zum Antiferromagnetismus (↑ 8.7.3.) beruht die Antiferroelektrizität auf der antiparallelen Ausrichtung der spontanen Polarisation, und zwar in einer Weise, daß sich die Feldwirkung der polaren Untergitter gegenseitig genau kompensiert. Diese Eigenschaft hat auch die Entdeckung dieser Erscheinung erschwert, denn erst die sehr gründliche Untersuchung eines Festkörpers erlaubt es, diese zunächst verdeckte Eigenschaft freizulegen. Wie in Abb. 8.6.4-6 skizziert ist, muß man sich eine antiferroelektrische Phase aus zwei ferroelektrischen Untergittern aufgebaut vorstellen. Makroskopisch verrät sich dieser Polarisationszustand weder durch ein elektrisches Feld an der Oberfläche des Festkörpers, noch durch eine Domänenstruktur. Erst anhand bestimmter Phasenumwandlungen und der damit verbundenen Anomalien kann man diesen Zustand aufspüren. In starken elektrischen Feldern ist ein Aufbrechen der antiparallelen Orientierung benachbarter Dipole möglich, und es entsteht erzwungenermaßen ein ferroelektrischer Zustand. Auch bei einem Antiferroelektrikum wird bei einem vollständigen Umlauf von einer hohen positiven zu einer hohen negativen Feldstärke und zurück eine bestimmte Hysteresekurve der Abhängigkeit $P(|E|)$ durchlaufen. Dabei setzen die Hystereseerscheinungen erst bei einer Feldstärke ein, bei der sich der ferroelektrische Zustand herausbildet (Abb. 8.6.4-7).

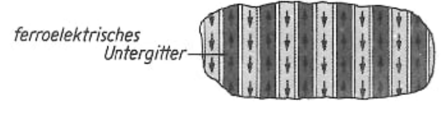


Abb. 8.6.4-6 Struktur eines Antiferroelektrikums

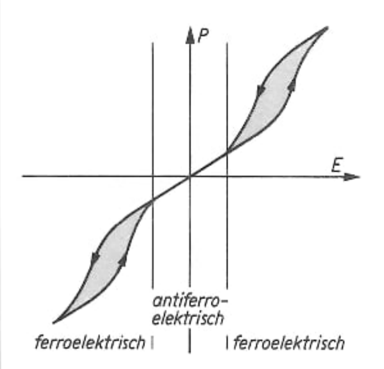


Abb. 8.6.4-7 Hysteresekurve eines Antiferroelektrikums

Von den mehr als 100 antiferroelektrischen Kristallen sollen als wichtige Beispiele Bleizirconat ($PbZrO_3$), Natriumniobat ($NaNbO_3$), Bleihafnat ($PbHfO_3$), Ammoniumdihydrogenphosphat ($NH_4H_2PO_4$) und Diammoniumtrihydrogenphosphat [$(NH_4)_2H_3IO_6$] genannt werden.

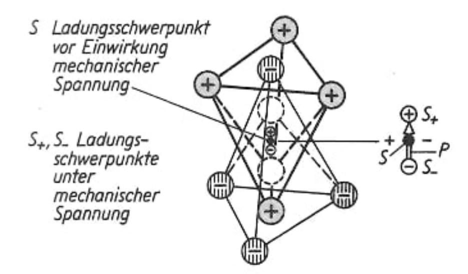


Abb. 8.6.5-1 Entstehung eines Dipolmoments durch mechanische Spannung in einem Piezoelektrikum

Auch für Antiferroelektrika lassen sich einige charakteristische Merkmale zusammenstellen:
Antiferroelektrika besitzen eine Curietemperatur. Oberhalb dieser sind sie parelektrisch. Unterhalb der Curietemperatur besitzen die Festkörper die niedrigere Kristallsymmetrie. In der Nähe des Curiepunkts ist das Curie-Weiß-Gesetz erfüllt. Durch Anwendung hoher Feldstärken kann der antiferroelektrische Zustand in den ferroelektrischen umgewandelt werden.
Antiferroelektrische Stoffe werden als Materialien für Speicherkondensatoren und elektroakustische Wandler verwendet. Für die weitere Entwicklung der verschiedensten Zweige der Festkörperelektronik ist es wesentlich, daß inzwischen auch günstige Eigenschaftskombinationen z. B. zwischen ferroelektrischen oder antiferroelektrischen Eigenschaften und magnetischen, halbleitenden oder supraleitenden Eigenschaften gefunden werden konnten.

8.6.5. Piezoelektrizität – Pyroelektrizität – Elektrostriktion

8.6.5.1. Piezoelektrizität. Eine Polarisation des Festkörpers kann bei einigen Klassen von Ionenkristallen auch durch Anlegen einer mechanischen Spannung erzeugt werden. Es handelt sich auch hier um die geeignete Schwerpunktverschiebung positiver und negativer Ladungen eines bestimmten Volumengebiets des Festkörpers, beispielsweise einer Elementarzelle, so daß ein Dipolmoment entsteht, wie aus Abb. 8.6.5-1 ersichtlich. Drückt man das durch elastische Kräfte zusammengehalten gedachte

Tetraeder auf die Spitze, so verschiebt sich sein Ladungsschwerpunkt S um ein kleines Stück nach unten. Werden zwei derartige Tetraeder, die jeder aus Bausteinen entgegengesetzter Polarität aufgebaut sind, in der skizzierten Weise kombiniert, so entsteht bei einer Kontraktion in der angegebenen Form ein Dipolmoment. Für nicht zu große mechanische Spannungen ist der Zusammenhang zwischen Polarisation und Dehnung und damit auch zwischen Polarisation und mechanischer Spannung linear. Piezoelektrische Kristalle besitzen einen von Null verschiedenen Piezomodul. Normalerweise sind die Zusammenhänge nur mit relativ komplizierten Tensorgleichungen zu erfassen. Wenn man aber polykristalline und damit isotrope *Piezokeramiken* beschreiben will, dann läßt sich die Abhängigkeit der Polarisation von einer mechanischen Druckspannung (z. B. auf die gegenüberliegenden Flächen einer Keramikplatte ausgeübt) durch $P = d\sigma$ ausdrücken. Hierin ist σ die mechanische Spannung und d der *Piezomodul*. Das ist die einfachste Beschreibung des *Piezoeffekts*. Eine Polarisation in einem Dielektrikum entsteht vor allem durch Anlegen eines äußeren Felds. Bei einem piezoelektrischen Festkörper muß demzufolge bei Anlegen einer elektrischen Feldstärke eine Dehnung entstehen. In stark vereinfachter Form kann das durch $\sigma = (1/d)P = (1/d)(\varepsilon - 1)\,\varepsilon_0\,|\boldsymbol{E}|$ beschrieben werden.

Von den 32 Kristallklassen erlauben 20 das Auftreten piezoelektrischer Strukturen. Voraussetzung ist, daß kein Symmetriezentrum im Gitter vorhanden ist (↑ 6.4.1.). Ferroelektrika sind in der polaren Phase grundsätzlich piezoelektrisch, aber nicht jeder piezoelektrische Stoff muß auch ferroelektrisch sein.

Technische Bedeutung hat vor allem Quarz. Da man durch genaue Untersuchungen weiß, in welchen kristallographischen Richtungen der größte Piezomodul auftritt, schneidet man ganz bestimmt orientierte Quarzstücke aus einem Einkristall heraus und versieht sie in der richtigen Feldrichtung mit Elektroden. Derartige Bauelemente bezeichnet man als Schwingquarze. Sie haben eine bestimmte Eigenfrequenz, die sehr konstant ist. Gewöhnlich wird diese Frequenzkonstanz mit $\Delta\nu/\nu = 10^{-8}$ angegeben. *Schwingquarze* oder *Quarzschwinger* sind in Quarzuhren, Quarzsendern, Frequenzmessern u. a. enthalten. Man kann sich die Frequenzstabilität anhand der Ganggenauigkeit einer *Quarzuhr* verdeutlichen. Eine durch einen Schwingquarz stabilisierte Zeitmessung weist auf 10^8 s einen Zählfehler von 1 s auf. Eine solche Uhr kann bei Einhaltung einer bestimmten Temperaturkonstanz nur 1 s in drei Jahren falsch gehen. Rechnet man mit den im täglichen Betrieb auftretenden Temperaturschwankungen, so reduziert sich die Ganggenauigkeit auf etwa ± 10 s pro Jahr. Aber auch dieser Wert gestattet es, eine solche Armbanduhr für ein ganzes Jahr mit Hilfe eines Uhrenschaltkreises zu programmieren, so daß eine völlig wartungsfreie exakte Zeitangabe möglich ist. Auch für die Messung der Schichtdicke kleinster abgeschiedener Stoffmengen haben solche Quarzschwinger große Bedeutung erlangt. Man kann nämlich die Verstimmung der Eigenfrequenz eines solchen Bauelements als Meßgröße für die Substanzmenge verwenden, die auf seiner Oberfläche aufgebracht wird. Diese Meßmethode ist so empfindlich, daß man schon 1/10 einer atomaren Bedeckung feststellen kann, andererseits aber auch in der Lage ist, das Wachstum einer dünnen Schicht dynamisch zu erfassen.

Für verschiedene Anwendungsgebiete wurden auch Piezokeramiken entwickelt. Aufgrund ihrer polykristallinen Struktur sind sie in größeren Volumenbereichen isotrop. Einige Beispiele zeigt dazu Tab. 8.6.5-2. Es genügt hier, den longitudinalen und den transversalen Piezomodul zu kennen. Zum Vergleich ist in der Tabelle einkristallines Bariumtitanat aufgeführt. Erst die Entwicklung geeigneter piezoelektrischer Festkörper hat eine Anwendung der Akustoelektronik eröffnet. *Akustoelektronische Wandler* werden in der Technik vielfältig verwendet, insbesondere auf der Basis von Quarz und Lithiumniobat. Ein solcher Wandler besteht aus einer piezoelektrischen Substratplatte, die an der Oberfläche zwei Paare kammartig ineinandergreifender Metallelektroden trägt. Benachbarte Finger von zwei Kammelektroden haben einen Abstand, der der halben Wellenlänge der anzuregenden mechanischen Schwingung entspricht. Wird an die Elektrode ein elektrisches Wechselfeld angelegt, dann entsteht an der Substratoberfläche eine mechanische Schwingung, die sich längs der Oberfläche mit einer charakteristischen Geschwindigkeit ausbreitet. Da bekanntlich Schallgeschwindigkeiten in Festkörpern einen Wert von einigen 1000 $\mathrm{ms^{-1}}$ haben, kommt das mechanische Signal am zweiten Kammelektrodenpaar verzögert an. Bei einem Abstand von 1 cm liegt die Verzögerung in der Größenordnung von 10 µs.

Tab. 8.6.5-2 Piezomoduln von Piezokeramiken; ε Dielektrizitätskonstante, d Piezomodul in $\mathrm{mV^{-1}}$

Piezokeramik	$\varepsilon \cdot 10^3$	$d_{trans} \cdot 10^{12}$	$d_{long} \cdot 10^{12}$
$Ba(Ti_{0,95}Zr_{0,05})O_3$	1,4	−60	150
$(Ba_{0,80}Pb_{0,08}Ca_{0,12})TiO_3$	0,6	−35	90
$(Na_{0,8}Cd_{0,2})NbO_3$	2,0	−80	200
$(Na_{0,75}Cd_{0,25})NbO_3$	2,0	−60	150
$(Pb_{0,7}Ba_{0,3})Nb_2O_6$	0,9	−40	100
$(Pb_{0,6}Ba_{0,4})Nb_2O_6$	1,5	−90	220

Da an der zweiten Elektrode aufgrund der mechanischen Deformation des Piezoelektrikums wieder ein elektrisches Wechselfeld entsteht, verzögert ein derartiger Wandler ein elektrisches Signal in definierter Weise. Gegenüber den bisherigen Formen von Verzögerungsgliedern ist das eine sehr elegante und ökonomische Lösung. Ein solcher Wandler hat gleichzeitig die Eigenschaften eines Filters, da nur ganz bestimmte mechanische Eigenfrequenzen angeregt werden können. Wird als Substratplatte ein halbleitendes Piezoelektrikum verwendet, dann können durch einen Elektronenstrom längs der Oberfläche die mechanischen Schwingungen sogar verstärkt werden. Auf diese Weise erhält man auch eine Verstärkung des verzögerten Wechselspannungssignals.

8.6.5.2. Elektrostriktion. Während beim Piezoeffekt über den Piezomodul eine vorzeichenbehaftete Beziehung zwischen einem an den Festkörper angelegten Feld und der mechanischen Deformation besteht – beide Fälle der Dehnung oder Kontraktion sind vorhanden –, tritt Elektrostriktion mehr oder weniger deutlich in jedem Festkörper auf.

Unter Elektrostriktion versteht man die Deformation eines Festkörpers unter der Wirkung eines äußeren elektrischen Felds, wobei dieser Effekt dem Quadrat der anliegenden Feldstärke proportional ist.

Die Deformation ist von der Feldstärkerichtung unabhängig. Vernachlässigt man alle Anisotropien, dann kann der Zusammenhang zwischen Dehnung $e = \Delta x/x_0$ und Feldstärke durch $e = qP^2 = q\varepsilon_0^2(\varepsilon - 1)^2 |E|^2$ dargestellt werden. Der *Elektrostriktionskoeffizient* q nimmt für Ferroelektrika Werte der Größenordnung 0,05 $m^4A^{-2}s^{-2}$ an. Kristalle, die einen hohen q-Wert besitzen, werden als Ultraschallgeber verwendet.

8.6.5.3. Pyroelektrizität. Nicht jedes Ferroelektrikum zeigt trotz im Inneren vorhandener spontaner Polarisation auch außen an der Oberfläche ein elektrisches Feld, abgesehen davon, daß das auch durch eine bestimmte Domänenanordnung erreichbar ist. Das sonst vorhandene elektrische Feld kann nämlich mittels Doppelschichten, die durch aus der Umgebung eingefangene Ladungsträger entstehen, neutralisiert werden. Die Entstehung einer derartigen Doppelschicht erfordert jedoch relativ viel Zeit. Wenn nun ein solcher Kristall erwärmt oder abgekühlt wird, ändert seine spontane Polarisation sehr schnell ihren Absolutwert. Für eine geänderte spontane Polarisation ist aber die momentan vorhandene Oberflächenladung nicht angepaßt. Als Nettoeffekt wird ein elektrisches Feld an der Oberfläche beobachtet. Deshalb kann man *pyroelektrische Kristalle* als Sonde für die Beobachtung von Wärmestrahlungen verwenden. Verstärkt wird diese Erscheinung noch dadurch, daß der durch die Wärmeausdehnung auftretende Piezoeffekt gleichsinnig wirkt.

8.7. Magnetische Erscheinungen in Festkörpern

8.7.1. Diamagnetismus

8.7.1.1. Verhalten von Stoffen im Magnetfeld. Schon in der Antike waren magnetische Kraftwirkungen bekannt, die natürlich vorkommende Eisenerze, z. B. Magnetit, aufeinander und auf Eisenstücke ausüben. 1820 entdeckte OERSTED die Wirkung des elektrischen Stroms auf einen Magneten. Später äußerte AMPERE den Gedanken, daß alle magnetischen Felder von elektrischen Strömen hervorgebracht werden. Diese Hypothese wurde Jahrzehnte danach durch die Entdeckung des magnetischen Moments von Elektronen und anderen Elementarteilchen glänzend bestätigt.

Obwohl es vor allem in der phänomenologischen Beschreibung von elektrischen und magnetischen Feldern eine große Ähnlichkeit gibt, führt gerade die nähere Betrachtung wesentlicher Unterschiede an die Problematik heran. Der Zusammenhang zwischen der magnetischen Induktion $\boldsymbol{B}$ und der magnetischen Feldstärke $\boldsymbol{H}$ wird für einen beliebigen Stoff durch die Beziehung $\boldsymbol{B} = \mu\mu_0\boldsymbol{H}$ beschrieben. (Permeabilität im Vakuum $\mu_0 = 1{,}2566 \cdot 10^{-6}$ $Vs(Am)^{-1}$). Die sogenannte *relative Permeabilität* μ ist eine Stoffkonstante und besitzt im allgemeinsten Fall die Eigenschaften eines Tensors. Für die meisten praktischen Fälle kann μ als einfache Zahl verwendet werden. Führt man analog zur Einführung der Polarisation (↑ 8.6.2.) eine Aufspaltung der magnetischen Induktion in einem Festkörper in zwei Anteile $\boldsymbol{B} = \mu_0\boldsymbol{H} + \boldsymbol{M} = \mu_0\boldsymbol{H} + \varkappa_{mag}\,\mu_0\boldsymbol{H}$ durch, so hat man den Festkörpereinfluß abgetrennt. Dieser wird durch die *Magnetisierung* $\boldsymbol{M}$ beschrieben. Relative Permeabilität μ und *magnetische Suszeptibilität* $\varkappa_{mag}$ hängen über $\mu = 1 + \varkappa_{mag}$ zusammen.

Tab. 8.7.1-1 Magnetische Suszeptibilität $\varkappa_{mag}$ einiger Stoffe; bei Eisen und Nickel hängen die gegebenen Werte von der Reinheit und der Vorbehandlung ab

Stoff	$\varkappa_{mag}$
Wasser	−0,0000091
Bismut	−0,00017
Sauerstoff	+0,0000019
Platin	+0,00027
Mangan	+0,00073
Eisen und Nickel (Anfangssuszeptibilität)	+100 bis 500

Wesentliche Unterschiede zu den Wirkungen elektrischer Felder bestehen auf folgenden Gebieten:
Die magnetische Suszeptibilität ist eine Größe, die im Unterschied zur elektrischen Suszeptibilität auch im statischen Fall sowohl positive als auch negative Werte annehmen kann. Anhand einer Messung von $\varkappa_{mag}$ läßt sich sofort entscheiden, ob ein Stoff diamagnetisch (−) oder paramangetisch (+) ist (Tab. 8.7.1-1).
Die Wechselwirkungsenergien atomarer magnetischer Momente sind um einen Faktor 10^{-5} kleiner als bei elektrischen Dipolmomenten.
Da in diamagnetischen und paramagnetischen Stoffen $|\boldsymbol{M}| \ll \mu_0 \boldsymbol{H}$ ist, kann für viele Berechnungen das in einem Festkörper entstehende zusätzliche magnetische Feld vernachlässigt werden. Das heißt aber, man kommt auch zu guten Ergebnissen, wenn für diesen Fall die Wechselwirkung magnetischer Dipole völlig vernachlässigt und auf die Berechnungsmethoden für isolierte magnetische Dipole, wie sie zum Beispiel auf Gase zutreffen, zurückgegriffen wird.
Elektrische und magnetische Felder haben ganz unterschiedliche Kraftwirkungen auf geladene Teilchen: $\boldsymbol{F}_{el} = q\boldsymbol{E}$ und $\boldsymbol{F}_{mag} = q\boldsymbol{v} \times \boldsymbol{B}$. Die Unterschiede bestehen in der Richtung und in der Geschwindigkeitsabhängigkeit.
Im Unterschied zu den elektrischen Eigenschaften der Stoffe kennt man beim Magnetismus bisher keine magnetischen Monopole oder magnetische Ladungen. Diese treten nur paarweise auf, z. B. bei einer Stromschleife oder bei einem Stabmagneten. Als Festlegung gilt, daß das Dipolmoment $\boldsymbol{m}_m$ vom Südpol zum Nordpol weist. Der Nordpol eines Dipols zeigt im Erdfeld auf den magnetischen Nordpol. Der Zusammenhang zwischen Magnetisierung und magnetischem Dipolmoment hat im einfachsten Fall die Form $\boldsymbol{M} = \mu_0 \sum \boldsymbol{m}_{m,i}$. Die Summation ist so zu führen, daß dadurch das resultierende magnetische Dipolmoment der Volumeneinheit ermittelt wird.

8.7.1.2. Magnetische Momente der atomaren Bausteine. In Abwesenheit eines magnetischen Felds sind die meisten Stoffe unmagnetisch. Oft wird erst in einem Magnetfeld innerhalb der atomaren oder molekularen Bausteine eines Festkörpers ein magnetisches Moment induziert. Warum nicht immer nach außen wirksame, permanente magnetische Momente vorhanden sind, ist nicht sofort einleuchtend, denn aus der Elementarteilchenphysik (↑ 3.2.) ist bekannt, daß sowohl das Elektron als auch Proton und Neutron ein permanentes magnetisches Moment besitzen. Außerdem erzeugt eine auf einer geschlossenen Bahn umlaufende Ladung nach

(1) $\boldsymbol{m} = q\boldsymbol{l}/2m$

ein magnetisches Moment. Hier ist $\boldsymbol{l}$ der Drehimpuls der Bahnbewegung und q und m sind Ladung und Masse des Teilchens. In diesem Abschnitt sollen einige Besonderheiten des Bahnanteils und des Spinanteils atomarer magnetischer Momente herausgearbeitet werden.
Ein Elektron, das einen Atomrumpf mit dem Bahndrehimpuls $\boldsymbol{l}$ umläuft, erzeugt ein magnetisches Moment $\boldsymbol{m} = -e\boldsymbol{l}/2m_0$. Da sich immer für ein Atom, Ion oder Molekül ein Gesamtbahndrehimpuls $\boldsymbol{L}$ angeben läßt, folgt sofort für den Gesamtbetrag des magnetischen Moments $\boldsymbol{m}_{ges} = -e\boldsymbol{L}/2m_0$. Nach der Quantentheorie lassen sich die atomaren Drehimpulse ihrem Betrag nach durch die Bahnimpulsquantenzahlen l bzw. L und die atomare Drehimpulseinheit $\hbar = h/2\pi$ ausdrücken: $|\boldsymbol{l}| = l\hbar$, $|\boldsymbol{L}| = L\hbar$. In einem Magnetfeld H_z stellen sich die atomaren oder molekularen magnetischen Dipole gequantelt ein. Das erfolgt so, daß $\Delta m_z = e\hbar/2m_0$ gilt. Es gibt also immer $2l + 1l_z$ bzw. m_z-Werte. Die Größe $e\hbar/2m_0$ wird als *Bohrsches Magneton* m_B bezeichnet und ist eine sehr gut angepaßte Einheit für die Beschreibung der Stärke atomarer magnetischer Dipole. Auf der Grundlage dieser Gesetzmäßigkeiten ist verständlich, warum viele Atome ein resultierendes magnetisches Moment Null besitzen. Für abgeschlossene Elektronenschalen und auch für abgeschlossene Unterschalen z. B. p-, d-, f-Schalen ist immer der Gesamtdrehimpuls $L = 0$. Einen Beitrag zum magnetischen Gesamtmoment $\boldsymbol{m}_{ges}$ leisten nur die Bahnmomente der Elektronen der nicht abgeschlossenen äußeren und inneren Schalen und Unterschalen eines Atoms. Dabei ist aber die wesentliche Einschränkung zu beachten, daß s-Elektronen keinen Bahndrehimpuls besitzen und damit auch kein *magnetisches Bahnmoment* haben können. Wasserstoff- oder Alkalimetallatome haben z. B. kein magnetisches Bahnmoment, weil sich die einzelnen äußeren Elektronen in s-Zuständen befinden. Bor- oder Aluminiumatome besitzen dagegen ein magnetisches Bahnmoment von der Größe eines Bohrschen Magnetons, weil sich das 3. Valenzelektron in einem p-Zustand mit $l = 1$ befindet.
Da ein wesentliches Prinzip für die Bindung von Atomen zu Molekülen und für die heteropolare und kovalente Bindung von Atomen zu Festkörpern darin besteht, daß die Valenzelektronen abgeschlossene Elektronenschalen bilden, gehören Festkörper im wesentlichen zu den Diamagnetika.
Auch die Eigenrotation der Elektronen, der Elektronenspin, ist mit einem magnetischen Moment verbunden. Allerdings ist hier eine andere Verknüpfung zwischen Eigendrehimpuls $\boldsymbol{s}$ und magnetischem Dipolmoment $\boldsymbol{m}_e = -\boldsymbol{s}e/m_0$ vorhanden. Da bekanntlich $|\boldsymbol{s}| = (1/2)\hbar$ gilt, ist $\boldsymbol{m}_e = e\hbar/2m_0$. Die fehlende 2 im Nenner wird kompensiert und man erhält wieder das Bohrsche Magneton. Auch Elektronenspins setzen

sich entsprechend der Russel-Saunders-Kopplung (↑ 4.2.4.4.) in einem Atom zu einem Gesamtspin S zusammen. Für einen atomaren oder molekularen Baustein ist ein *magnetisches Gesamtspinmoment* von $|m_s| = Se\hbar/2m_0$ zu erwarten. Ein Elektron leistet zum Gesamtmoment einen Beitrag durch seinen Elektronenspin wie er auch durch eine Bahnbewegung im p-Zustand $l = 1$ auftreten würde. Da aber das magnetische Spinmoment des Elektrons um 0,116% größer ist als das Bohrsche Magneton, ist diese Übereinstimmung nur näherungsweise vorhanden. Auch die magnetischen Gesamtspinmomente von abgeschlossenen Schalen und Unterschalen verschwinden. Bei der Bildung von Molekülen und von Festkörpern aus Atomen kommt es zu einer weitgehenden Kompensation vorher in den atomaren Bausteinen vorhandener magnetischer Spinmomente. Das ist z. B. für das Wasserstoffmolekül oder das Alkali-Halogenid-Gitter (NaCl, KBr u. a.) der Fall. Während ein isoliertes Natrium- oder Chloratom ein resultierendes von Null verschiedenes *magnetisches Spinmoment* besitzt, ist das für ein Na^+- oder Cl^--Ion nicht mehr der Fall, weil sich hier abgeschlossene Elektronenschalen ausbilden. Beim Wasserstoffmolekül existiert auch der höher angeregte Zustand der Parallelorientierung der beiden beteiligten Elektronenspins. Der Grundzustand, der für flüssigen oder festen Wasserstoff charakteristisch ist, ist aber immer durch die Antiparallelstellung der Elektronenspins gekennzeichnet. Das ist typisch für den Aufbau aller Gasmoleküle oder aller organischen Moleküle. Für die aus solchen Bausteinen aufgebauten Molekülkristalle vom Van-der-Waals-Typ wird man deshalb diamagnetisches Verhalten erwarten.

Schließlich bleibt noch der Beitrag des magnetischen Moments der Atomkerne zum magnetischen Verhalten zu betrachten. Die *magnetischen Kernmomente* (↑ 3.2.1.) liegen in der Größenordnung des Kernmagnetons $m_K = e\hbar/m_p$, m_p ist die Masse des Protons. Sie sind aber keine ganzzahligen Vielfachen dieses Werts, da schon die einfachsten Kerne wie Proton und Neutron vom Kernmagneton abweichende Werte besitzen. Aufgrund der Relation $m_p/m_0 = 1836$ sind die magnetischen Kernmomente um einen Faktor von ungefähr 10^3 kleiner als die durch Elektronen hervorgerufenen. Für die Entscheidung Diamagnetismus oder Paramagnetismus fallen die Kernmomente deshalb nicht ins Gewicht.

8.7.1.3 Magnetisches Verhalten und seine Deutung. In diamagnetischen Stoffen ist die magnetische Suszeptibilität negativ und damit die Magnetisierung dem von außen angelegten Feld entgegengerichtet. Nach dem Induktionsgesetz wird in einem Stoff, wenn ein Magnetfeld zugeschaltet wird, durch die Änderung des Induktionsflusses eine Spannung induziert, die ihrerseits Kreisströme bewirkt. Diese *Kreisströme* kann man sich zusammengesetzt denken aus vielen mikroskopisch kleinen Kreisströmen, die in atomaren bzw. molekularen Bausteinen fließen. Die Kreisströme fließen völlig reibungsfrei und bleiben, einmal angeworfen, auch bei konstanter magnetischer Feldstärke bestehen. Auf diese Weise werden in jedem Baustein induzierte magnetische Momente hervorgerufen, die nach der Lenzschen Regel dem den Vorgang auslösenden Magnetfeld entgegengerichtet sind.

Da die induktive Beeinflussung der Elektronenbewegung für jeden Stoff vorhanden ist, ist im Prinzip auch jeder Stoff diamagnetisch. Sind nun aber permanente magnetische Dipole aufgrund des Bahn- oder Spinmagnetismus vorhanden, wird der *Diamagnetismus* durch den Paramagnetismus oder Ferromagnetismus überdeckt.

Diamagnetisch sind die meisten Gase, anorganische oder organische Isolierstoffe, viele wichtige Halbleiter mit kovalenter Bindung vom Element- oder Verbindungs-Halbleitertyp, Ionenkristalle und auch einige Metalle wie Kupfer und Bismut.

Die magnetische Suszeptibilität der *Diamagnetika* ist bei chemisch ähnlich gebauten Stoffen um so größer, je mehr Elektronen die atomaren oder molekularen Bausteine des betreffenden Stoffs enthalten. Bei den oft sehr asymmetrisch aufgebauten organischen Molekülkristallen, z. B. Naphthalin oder Anthrazen, ist die Suszeptibilität eine stark anisotrope Größe. Zur exakten Beschreibung des magnetischen Verhaltens derartiger Molekülkristalle muß $\varkappa_m$ als Tensorgröße formuliert werden. Aufgrund der im atomaren Bereich liegenden Wirkmechanismen ist die diamagnetische Suszeptibilität, wie auch die Beschaffenheit der Elektronenhülle, nur sehr schwach temperaturabhängig.

Die klassische Theorie des Diamagnetismus wurde von LANGEVIN entwickelt. Ausgangspunkt ist dabei das von LARMOR aufgestellte Theorem, daß sich der unveränderten Bewegung der Hüllelektronen durch Zuschalten eines Ma-

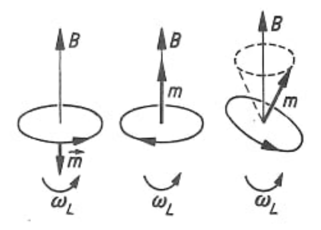


Abb. 8.7.1-2 Larmorpräzession ω umlaufender Elektronen; $\boldsymbol{m}$ Dipolmoment, $\boldsymbol{B}$ äußeres Feld

gnetfelds in erster Näherung eine Präzessionsbewegung überlagert. Die Kreisfrequenz der *Larmorpräzession* ist $\omega_L = eB/2m_0$. B ist die magnetische Induktion. Man erhält den Wert der Larmorfrequenz durch Aufsuchen der Gleichgewichtsbedingung zwischen Zentrifugalkraft und Lorentzkraft (Abb. 8.7.1-2). Obwohl die richtige Formel für die magnetische Suszeptibilität sowohl klassisch als auch quantenmechanisch abgeleitet werden kann, soll hier kurz der quantenmechanische Berechnungsweg angedeutet werden. Ein einzelnes Elektron besitzt in einem äußeren Magnetfeld H_z die durch den Hamiltonoperator beschriebenen Energieanteile: die kinetische Energie $-(\hbar^2/2m_0)(\partial^2/\partial r^2)\psi$, die potentielle Energie E_p (in der die Elektron-Kern- und die Elektron-Elektron-Wechselwirkung sowie den betreffenden Elektronenzustand charakterisierende Energien enthalten sein sollen), den Energieanteil der Präzession der vorhandenen Bahn- und Eigendrehimpulse $\omega_L(l_z + 2s_z)$, die Spin-Bahn-Wechselwirkungsenergie E_{SL} und das für diese Überlegungen wichtige Glied der Rotationsenergie des Elektrons infolge der Präzessionsbewegung $(1/2) \times m_0\omega_L^2\bar{r}^2$. Mit $\bar{r}^2$ ist hier das Quadrat des mittleren Abstands einzusetzen, den ein Elektron von der durch den Atomkern gehenden Achse der Präzessionsbewegung hat. Da aus Messungen bekannt ist, daß $\varkappa_{mag}$ eine sehr kleine Größe ist, kann der für den Diamagnetismus verantwortliche Energieanteil als eine kleine Korrekturgröße aufgefaßt werden. Für ein Ensemble von Elektronen kann man die Energie deshalb bei Anliegen eines Magnetfelds in der folgenden Form aufschreiben:

$$(2) \quad E = E_1 + E_2.$$

E_1 umfaßt die Summe aller hier nicht weiter interessierenden Energieanteile, auch über alle beteiligten Elektronen summiert, und E_2 ist

$$(3) \quad E_2 = (1/2)\, m_0\omega_L^2 \sum_n \bar{r}_n^2.$$

Da die Energie, die ein magnetisches Dipolmoment in einem Magnetfeld besitzt

$$(4) \quad E = -\boldsymbol{m}\boldsymbol{B} = -m_z B_z$$

ist, ergibt sich das induzierte Dipolmoment aus 8.7.(3) und 8.7.(4) zu

$$(5) \quad m_z = -\mathrm{d}E_2/\mathrm{d}B_z = -m_0(eB_z/2m_0)\,(e/2m_0) \sum_n \bar{r}_n^2 = -(e^2/4m_0)\, B_z \sum_n \bar{r}_n^2.$$

Aus 8.7.(5) kann mit Hilfe der in Abschnitt 8.7.1.1. gegebenen Definition für die Suszeptibilität weitergerechnet werden, um $\varkappa_{mag}$ zu erhalten. Es gilt

$$(6) \quad \varkappa_{mag} = M/\mu_0 H_z = \mu_0 N m_z/B_z.$$

N ist die Zahl der Atome bzw. Moleküle pro Volumeneinheit. Die relative Permeabilität μ wurde in dieser Näherung $\mu = 1$ gesetzt. Da es nicht um die Berechnung von Anisotropien geht, gilt die Annahme, daß die atomaren Bausteine bezüglich der Richtung des magnetischen Felds statistisch orientiert sind, und für jedes Glied kann $\bar{r}_n^2 = (2/3)r_{n_0}^2$ gesetzt werden. Die Größe $r_{n_0}^2$ ist dabei das Quadrat des mittleren Abstands eines Hüllelektrons vom Kern. Man findet also schließlich:

$$(7) \quad \varkappa_{mag} = -(Ne^2\mu_0/6m_0) \sum_n r_{n_0}^2.$$

Die aus dieser Formel berechneten Werte, wobei die Erwartungswerte der Aufenthaltswahrscheinlichkeit der Elektronen zugrunde gelegt werden können, ergeben eine recht gute Übereinstimmung mit dem Experiment. Bei der Berechnung der magnetischen Eigenschaften von Metallen und hochdotierten Halbleitern muß berücksichtigt werden, daß der Diamagnetismus der Gitterbausteine mit dem Paramagnetismus der Leitungselektronen in Konkurrenz tritt. Da beide Anteile etwa von derselben Größenordnung sind, können diese Festkörper sowohl Diamagnetika als auch Paramagnetika sein.

8.7.2. Paramagnetismus

8.7.2.1. Paramagnetische Stoffe. Paramagnetisch verhalten sich Stoffe, deren Gesamtbahn- und Spinmomente die Entstehung eines *permanenten magnetischen Moments* zulassen. Die magnetischen Kernmomente reichen dagegen nicht aus, um den immer vorhandenen Diamagnetismus zu überdecken.
Die wichtigsten Klassen paramagnetischer Stoffe sind:

- Alle Stoffe, die als endgültige Bausteine eine ungerade Zahl von Elektronen haben, z. B. Alkalimetalldampf, Stickstoffmonoxid, freie organische Radikale; diese Bedingung führt auf einen von Null verschiedenen Gesamtspin.
- Stoffe, deren Atome oder Ionen teilweise gefüllte innere Elektronenschalen enthalten wie die Übergangselemente, die seltenen Erdmetalle und die Actiniden, z. B. die Ionen Ti^{3+}, Mn^{2+}, Fe^{3+}, Ni^{2+}, Ce^{3+} und U^{4+}. Die meisten dieser paramagnetischen Ionen behalten diese Eigenschaft auch in Lösungen oder wenn sie in beliebigen Festkörpern eingebaut sind. In der Analytik benutzt man dieses Verhalten für den Nachweis paramagnetischer Verunreinigungen mit Hilfe der Elektronenspinresonanz (↑ 4.4.2.).
- Verschiedene chemische Verbindungen mit gerader Elektronenzahl, aber mit einer speziellen Besetzung der Elektronenzustände, z. B. das Sauerstoffmolekül O_2.

- Viele Metalle, bei denen der Paramagnetismus der Leitungselektronen den Diamagnetismus übertrifft.

8.7.2.2. Theoretische Deutung des Paramagnetismus. Im allgemeinen tragen Spin- und Bahndrehimpuls zum magnetischen Moment bei. Es ist also vom Gesamtdrehimpuls des Bausteins auszugehen. Unter den Bedingungen, die hier interessieren, ist in einem Elektronensystem *Russel-Saunders-Kopplung* vorhanden. Magnetisches Moment und Gesamtdrehimpuls besitzen nicht die gleiche Richtung. Das stellt Abb. 8.7.2-1 übersichtlich dar. Da die Kopplung von

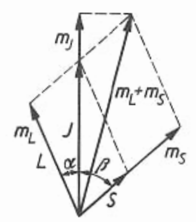


Abb. 8.7.2-1 Magnetisches Moment eines Gitterbausteins (Erläuterung im Text)

Spin- und Bahndrehimpuls des einzelnen Elektrons vernachlässigt werden kann, finden wir den Bahndrehimpuls als Summe der Bahndrehimpulse $\boldsymbol{l}_i$ der einzelnen Elektronen $\boldsymbol{L} = \sum \boldsymbol{l}_i$ und entsprechend auch den Gesamtspin $\boldsymbol{S} = \sum \boldsymbol{s}_i$. Daraus läßt sich der Gesamtdrehinpuls $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$ aufbauen. Sein quantenmechanischer Erwartungswert ist $|\boldsymbol{J}| = \hbar(J(J+1))^{1/2}$. Berücksichtigt man, daß das magnetische Bahnmoment aus $\boldsymbol{m}_L = -g_L m_B(\boldsymbol{L}/\hbar)$ mit $g_L = 1$ und das magnetische Spinmoment aus $\boldsymbol{m}_S = -g_S m_B(\boldsymbol{S}/\hbar)$ mit $g_S = 2{,}0023 \approx 2$ gewonnen werden kann, dann folgt unmittelbar die obige Feststellung.

Aufgrund des Erhaltungssatzes für den Drehimpuls ist $\boldsymbol{J}$ ohne äußere Einwirkung raumfest. Das resultierende magnetische Moment $\boldsymbol{m}_{res} = \boldsymbol{m}_L + \boldsymbol{m}_S$ präzessiert deshalb um die Richtung des Gesamtdrehimpulses, und nach außen ist nur seine Komponente m_J wirksam. Mit Hilfe des Kosinussatzes kann der für den Zusammenhang zwischen $\boldsymbol{m}_J$ und $\boldsymbol{J}$ wesentliche Landéfaktor g in $\boldsymbol{m}_J = g m_B(\boldsymbol{J}/\hbar)$ berechnet werden. Man hat anzusetzen:

$$(8) \qquad m_J = m_B(L\cos\alpha + 2\,S\cos\beta).$$

Daraus folgt:

$$m_J = \left(\frac{J^2 + L^2 - S^2}{2J} + \frac{2(J^2 + S^2 - L^2)}{2J)}\right) m_B$$

und nach Einsetzen der quantenmechanischen Erwartunswerte

$$(9) \qquad m_J = \frac{3J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} \times [J(J+1)]^{1/2}\, m_B = g\,|\boldsymbol{J}|\, m_B$$

Normalerweise sind die magnetischen Momente paramagnetischer Stoffe durch die Wärmebewegung regellos orientiert, so daß im Mittel kein resultierendes Moment vorhanden ist. In einem Magnetfeld sind aber die verschiedenen Einstellmöglichkeiten der magnetischen Dipole nicht mehr gleich wahrscheinlich. Es entsteht eine Vorzugsorientierung parallel zur Richtung des Magnetfelds. Allerdings können magnetische Dipole in einem Magnetfeld nur ganz bestimmte Lagen einnehmen. Die Einstellmöglichkeiten werden in Abb. 8.7.2-2 an Beispielen erläutert. Diese zugelassenen diskreten Lagen folgen der Festlegung, daß das magnetische Moment m_J bezüglich der Feldrichtung von $\boldsymbol{H}$ nur die Komponenten $m_{JH} = g m_B M_J$ haben kann. M_J ist die magnetische Orientierungsquantenzahl und durchläuft die Werte $M_J = +J, J-1, \ldots, -J$; $\Delta M_J = \pm 1$. Ein magnetisches Moment besitzt also in einem äußeren Magnetfeld $2J + 1$ Einstellmöglichkeiten. Ist reiner Spinmagnetismus vorhanden, gilt für die Zahl der Einstellmöglichkeiten $2S + 1$ bzw. für reinen Bahnanteil $2L + 1$. Mit Hilfe der Boltzmannstatistik läßt sich berechnen, in welcher Weise sich ein Gleichgewicht zwischen der ausrichtenden Wirkung des Magnetfelds und der den Ordnungsvorgang störenden Wärmebewegung zu erwarten ist. Für das Verhältnis von zum Feld parallel orientierten Dipolen $N_{\uparrow\uparrow}$ zu antiparallel orientierten Dipolen $N_{\uparrow\downarrow}$ erhält man $N_{\uparrow\uparrow}/N_{\uparrow\downarrow} = \exp(+2m_J B/k_B T)$. $2m_J B$ ist die Energiedifferenz zwischen den potentiellen Energien eines Dipols in Parallel- und Antiparallelstellung. Mit $m_J = m_B = 9{,}27 \cdot 10^{-24}\ \mathrm{Am^2}$ und dem sehr hohen Wert $B = \mu_0 \cdot 10^6\ \mathrm{Am^{-1}} = 1{,}26\ \mathrm{T}$ ergibt sich für diese Energiedifferenz ein Wert von $1{,}46 \cdot 10^{-4}$ eV. Bei Zimmertemperatur $kT = 0{,}025$ eV geht damit in die Exponentialfunktion ein Exponent von 0,00584 ein. Das ergibt für $N_{\uparrow\uparrow}/N_{\uparrow\downarrow}$ einen Wert von 1,006, also eine nur geringe Abweichung von der durch die Wärmebewegung erzwungenen statistischen Orientierung.

Ein magnetisches Feld bewirkt bei Zimmertemperatur in paramagnetischen Stoffen nur eine sehr kleine Vorzugsorientierung. Nur bei sehr

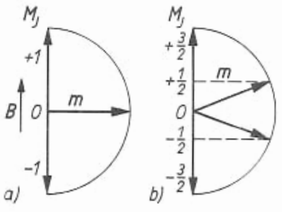


Abb. 8.7.2-2 Einstellmöglichkeiten der atomaren magnetischen Momente $\boldsymbol{m}$ zum äußeren Magnetfeld $\boldsymbol{B}$ für a) $L = 1$, $S = 1$, b) $S = 3/2$

tiefen Temperaturen gelingt es bei Anwendung sehr hoher Magnetfelder, die atomaren Dipole fast vollständig auszurichten und einen Sättigungswert der Magnetisierung zu erreichen. Das zu dieser Abschätzung verwendete Modell entspricht dem Einstellvorgang von Bausteinen mit dem Gesamtspin $S = 1/2$.

Rechnungen für komplizierter beschaffene magnetische Dipole sind schon sehr aufwendig. Deshalb geben wir hier nur das Endergebnis für die magnetische Suszeptibilität an, wenn die Temperatur noch nicht zu niedrig ist: $\varkappa_{mag} = \mu_0 N m_J^2 / 3kT$. Charakteristisch für den Paramagnetismus ist, daß die Suszeptibilität umgekehrt proportional mit der Temperatur abnimmt, wie es durch die zunehmende Störung der Dipolorientierung infolge der Wärmebewegung zu erwarten ist. Empirisch wurde dieser Zusammenhang schon vor der theoretischen Begründung von P. CURIE gefunden. Man bezeichnet die Beziehung $\varkappa = C/T$ deshalb als das Curiesche Gesetz. $C = \mu_0 N m_J^2 / 3k$ ist die Curiekonstante.

8.7.2.3. Paramagnetismus von Leitungselektronen. Wenn man den Beitrag der Leitungselektronen zur Suszeptibilität erfassen will, dann ist zu berücksichtigen, daß sie der Fermistatistik unterliegen. Das magnetische Moment von Leitungselektronen kann sich in einem Magnetfeld mit $2S + 1 = 2$ Einstellmöglichkeiten $+m_B$ oder $-m_B$ ausrichten. Da sich die beiden Einstellungen energetisch um $\Delta E = 2m_B B$ unterscheiden, wird die *Zustandsdichteverteilung der Elektronen eines Metalls in einem Magnetfeld* verändert. Wie das Abb. 8.7.2-3 deutlich macht, muß man dann zwei Zustandsdichteverläufe $D_{\uparrow\uparrow}(E)$ und $D_{\uparrow\downarrow}(E)$ unterscheiden. Da weiterhin die Maximalenergie der Ladungsträger durch die Fermienergie gegeben ist, entsteht ein Überschuß von Elektronen mit Parallelorientierung zum äußeren Feld. Für die Berechnung des Suszeptibilitätsanteils ist nur die Differenz der parallel und antiparallel orientierten Elektronenspins entscheidend. Man kann also ansetzen:

$$(10) \quad \varkappa_{mag} = [m_B(N_{\uparrow\uparrow} - N_{\uparrow\downarrow})]/H$$

Da die Fermienergie sehr viel größer ist als $2m_B B$, liegt hier wieder ein ähnlicher Fall vor, wie er schon bei der Behandlung der spezifischen Wärmekapazität der Ladungsträger eines Metalls auftrat (↑ 8.4.3.5.). Ohne einen großen Fehler zu machen, kann man davon ausgehen, daß von einer geringfügigen energetischen Umverteilung nur Elektronen in unmittelbarer Umgebung der Fermifläche betroffen werden (s. a. 8.4.4.2.). Für die Zahl der überschüssigen parallel orientierten Spins ergibt sich darum:

$$(11) \quad N_{\uparrow\uparrow} - N_{\uparrow\downarrow} = f(E_F) D(E_F)\, 2m_B B = m_B B (3n/2E_F).$$

Das letzte Gleichheitszeichen geht auf die Anwendung von Formel 8.4.(20) zurück. Die für Metalle und entartete Halbleiter wichtige *Paulische Suszeptibilität* der Leitungselektronen ergibt sich schließlich zu: $\varkappa_{Pauli} = 3n\mu_0 m_B^2 / 2E_F$. Die Größe n ist hier die Konzentration der freien Ladungsträger. In Tab. 8.7.2-4 werden einige Werte für die paramagnetische Suszeptibilität gegeben.

8.7.2.4. Erzeugung sehr tiefer Temperaturen durch adiabatische Entmagnetisierung. Bei tiefen Temperaturen kann durch ein starkes Magnetfeld eine nahezu vollständige Ausrichtung der permanenten Dipole eines Paramagnetikums erreicht werden. Diese Eigenschaft bildet eine der wesentlichen Grundlagen der Kryotechnik zur Erzeugung extrem tiefer Temperaturen. Man geht dabei so vor, daß ein geeignetes *paramagnetisches Salz*, z. B. Eisen-Ammonium-Alaun $[Fe_2(SO_4)_3(NH_4)_2SO_4 \cdot 24H_2O]$ in einem Heliumkryostaten auf 1 K vorgekühlt wird. Das erfolgt bekanntlich dadurch, daß man das flüssige Helium durch Druckerniedrigung zum Sieden bringt, wobei die Verdampfungswärme dem flüssigen Helium entzogen wird und eine Abkühlung des Heliums unter seinen Siedepunkt bei Normaldruck von 4,2 K ermöglicht wird. Mit einem Magnetfeld von mehr als 10^6 Am^{-1} wird das paramagnetische

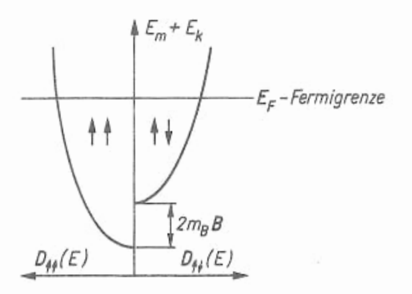


Abb. 8.7.2-3 Besetzung der Elektronenniveaus im Magnetfeld (Erläuterung im Text)

Tab. 8.7.2-4 Spezifische Suszeptibilitäten paramagnetischer Stoffe

Stoff	$\varkappa/\varrho$ in 10^{-5}cm^{3}g^{-1}
Al	0,77
Fe (1 200 °C)	33
Ni (800 °C)	15,1
Cr	3,7
U	0,52
W	0,30
$CoCl_2$	128
CuO	3,8
FeS	18

Salz dann in einem isothermen Vorgang ausgerichtet. Die bei der Ausrichtung frei werdende Ordnungsenergie $T\Delta S$ wird dabei vom Kältebad aufgenommen. Die Entropie S des Salzes wird bei diesem Ordnungsvorgang um den Betrag ΔS abgesenkt. Einen Kühlungseffekt erzielt man dadurch, daß nahezu jeder Wärmeaustausch zwischen Heliumkältebad und paramagnetischem Salz unterbunden wird. Das wird erreicht durch Einhaltung definierter Wärmestrahlungsbedingungen und durch Erzeugung eines Vakuummantels um das Salz herum. Wird dann unter diesen Bedingungen das Salz durch Abschalten des äußeren Magnetfelds adiabatisch entmagnetisiert, kühlt es sich, auch in einem mehrstufigen Prozeß, günstigstenfalls auf 1 mK ab. Nach Abschalten des Magnetfelds werden die atomaren Dipole nicht mehr gehindert, sich räumlich statistisch zu orientieren. Die atomaren Dipole führen dann für ihren thermodynamischen Zustand charakteristische Drehschwingungen aus. Die für diese Wärmebewegung erforderliche Energiemenge wird adiabatisch dem paramagnetischen Salz entzogen.

Zur quantitativen Beschreibung der adiabatischen Entmagnetisierung kann man ein Zustandsdiagramm $S(T)$ mit der magnetischen Feldstärke H als Parameter heranziehen. Die isotherme Magnetisierung $A \rightarrow B$ ist mit einer Abnahme der Entropie verbunden. Die $S(H, T)$-Kurven haben den in der Abb. 8.7.2-5 skizzierten Verlauf. Man erhält also je nach der herrschenden Feldstärke zwei Kurven $S(O, T)$ und $S(H, T)$. Der Versuchsverlauf $B \rightarrow C$ wird durch die adiabatische Entmagnetisierung erzwungen. Nach den Gesetzen der Thermodynamik erfolgt ein adiabatischer Vorgang immer so, daß die Entropie S dabei konstant bleibt. Also stellt sich nach Beendigung des Versuchs die Temperatur T_2 ein, d. h., die Entmagnetisierung führt parallel zur T-Achse zurück zur ersten Entropiekurve. Die in einer logarithmischen Temperaturskala besonders deutlich hervortretende relativ große Änderung der Temperatur kommt dadurch zustande, daß im betrachteten Temperaturbereich die üblichen Gitterschwinungen nur noch sehr schwach oder gar nicht mehr angeregt sind. Der Hauptteil der Wärmebewegung wird durch die Schwingung der Dipolmomente getragen, aber gerade dieser Anteil wird durch die Magnetisierung des Stoffs weitgehend eingeschränkt.

Die Temperaturgrenze von etwa 1 mK kann auf die beschriebene Weise zunächst infolge der Nullfeldaufspaltung nicht unterschritten werden. Unter *Nullfeldaufspaltung* versteht man die ohne äußeres Feld auftretende Parallelorientierung magnetischer Dipole. Nullfeldaufspaltung tritt vor allem bei sehr tiefen Temperaturen auf und geht auf die Dipol-Dipol-Wechselwirkung zurück. Um die Nullfeldaufspaltung weitgehend zu unterdrücken, gibt es prinzipiell zwei Möglichkeiten, entweder Verdünnen des paramagnetischen Salzes oder Verwenden von Stoffen, die nur einen Kernparamagnetismus aufweisen. Als günstige Variante hat sich für die Erzeugung von Temperaturen bis hinab zu 1 μK die Versuchsdurchführung nach Abb. 8.7.2-6 erwiesen. Man bringt zunächst auf die oben beschriebene Weise ein paramagnetisches Salz auf etwa 10 mK und erlaubt dem kernparamagnetischen Salz, sich über den Wärmeschalter, z. B. einen dünnen Bleidraht, auf diese Temperatur einzustellen. Bei dieser Temperatur magnetisiert man auch den zweiten Stoff isotherm. Dann werden durch Umschalten des Bleidrahts vom normalleitenden in den supraleitenden Zustand adiabatische Bedingungen eingestellt (↑ 8.5.2.3.). Das geschieht einfach durch Herausziehen des Kryostaten aus dem Magnetfeld. Gleichzeitig setzt dann die Entmagnetisierung ein. Vorteilhaft wird als zweiter Stoff Kupfer verwendet.

8.7.2.5. Paramagnetische Resonanzabsorption. Magnetische Dipole, z. B. magnetische Spinmomente oder magnetische Kernmomente stellen sich in einem äußeren Magnetfeld in diskreter Weise ein (↑ 8.7.2.3.). Als Auswahlregel für die Änderung einer Orientierung im Magnetfeld ist für die Orientierungsquantenzahl $\Delta M_J = \pm 1$ gefordert. Ändert sich nun die Orientierung eines elementaren Dipols im Magnetfeld, dann ist dazu die Energie $\Delta E = h\nu = g m_B B$ erforderlich. Ein solcher Umklappvorgang kann durch Einstrahlen einer Mikrowellen-

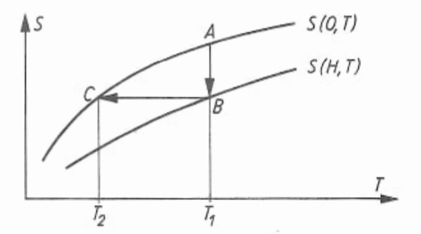


Abb. 8.7.2-5 Adiabatische Entmagnetisierung im Entropie-Temperatur-Diagramm (Erläuterung im Text)

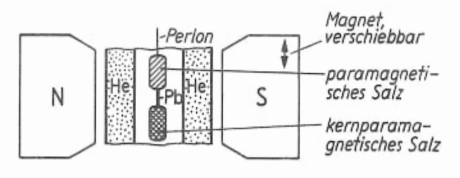


Abb. 8.7.2-6 Prinzip des Versuchsaufbaus zur Erzeugung extrem tiefer Temperaturen

strahlung der Quantenenergie $h\nu$ erzwungen werden. Bei diesem Vorgang wird dem Mikrowellenfeld Energie entzogen. Anhand des Absorptionsvorgangs lassen sich also die Energieniveaus auffinden, die für den Umklappvorgang der elementaren Dipole typisch sind (↑ 4.4.).

8.7.3. Ferromagnetika

8.7.3.1. Typische Eigenschaften der Ferromagnetika. Magnetwerkstoffe im engeren Sinne sind Stoffe, die eine spontane Magnetisierung besitzen. Während jeder Stoff unter bestimmten äußeren Bedingungen entweder diamagnetisch oder paramagnetisch ist, tritt eine *spontane Magnetisierung* nur bei einer sehr beschränkten Zahl von Stoffen auf. WARBURY (1880) und EWING (1882) entdeckten unabhängig voneinander die Hystereseerscheinung in Magnetwerkstoffen vom ferromagnetischen Typ. Seinen bedeutendsten Anstoß erhielt das Gebiet durch die Entdeckung von P. CURIE, daß der Ferromagnetismus oberhalb einer bestimmten Temperatur in Paramagnetismus übergeht. Die typischen Eigenschaften der Magnetwerkstoffe lassen sich in der folgenden Aufstellung zusammenfassen:

Magnetwerkstoffe besitzen eine spontane Magnetisierung. Diese ist auf bestimmte kristallographische Richtungen in einem einkristallinen Festkörper beschränkt.

Die spontane Magnetisierung ist bezüglich ihrer Richtung umkehrbar.

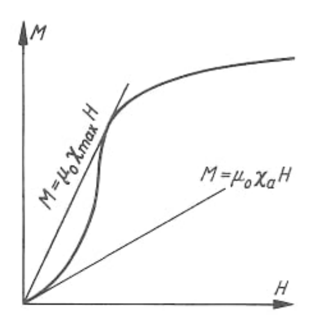


Abb. 8.7.3-1 Neukurve (Erläuterung im Text)

Bei einem vollständigen Durchlaufen des Umpolens der spontanen Magnetisierung von der Ausgangslage in die entgegengesetzte Richtung und zurück wird eine typische Hysteresekurve beobachtet.

Spontane Magnetisierung tritt in den meisten Fällen nicht bei allen Temperaturen auf, die für die Existenz eines Festkörpers verträglich sind. Oberhalb einer charakteristischen Temperatur liegen die meisten Magnetwerkstoffe in einer nichtpolaren Phase vor, wobei paramagnetische Eigenschaften vorherrschen.

In der Umgebung der Umwandlungstemperatur besitzt die relative Permeabilität ein ausgeprägtes Maximum mit Spitzenwerten der Größenordnung 10^6.

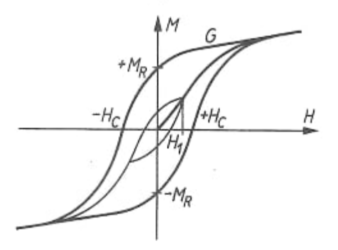


Abb. 8.7.3-2 M-H-Diagramm mit Neukurve, Hysteresekurve zur maximal angelegten Feldstärke H_1 und Grenzschleife G

Die Umwandlung von einer Phase in die andere ist mit einer Änderung der Ordnung der Bausteine im Nahordnungsbereich verbunden.

Ein spontan magnetisierter Kristall kann mit einheitlicher Magnetisierungsrichtung als Eindomänenkristall oder in Raumgebiete unterschiedlicher Magnetisierung eingeteilt auftreten.

Zu den Ferromagnetika gehören die Elemente Fe, Co, Ni, Gd, Tb, Dy, Ho, Er, Tm und viele Legierungen, z. B. Permalloy, die Heuslersche Legierung, Supermalloy oder Alnico.

8.7.3.2 Magnetisierungsvorgang bei Ferromagnetika. Stellt man einen ferromagnetischen Festkörper in einem völlig magnetfreien Raum aus der Schmelze her, dann zeigt er makroskopisch keine Magneteigenschaften. Im Magnetisierungs-Feldstärke-Diagramm [$M(H)$-Diagramm] kann für einen derartig hergestellten Festkörper die Neukurve (Abb. 8.7.3-1) gemessen werden. Aus dieser $M(H)$-Abhängigkeit läßt sich die magnetische Anfangssuszeptibilität $\varkappa_a$ und die maximale Suszeptibilität $\varkappa_{max}$ gewinnen. Bricht man irgendwo auf der Neukurve den Magnetisierungvorgang ab und reduziert von dieser Position aus die Stärke des Magnetfelds wieder, dann zeigt sich, daß der Magnetisierungsvorgang nicht reversibel abläuft. Einen vollständigen Umlauf mit getrennten Wegen für den Magnetisierungs- und Entmagnetisierungsvorgang, meist zwischen Werten, die schon im Sättigungsbereich der Magnetisierung liegen, bezeichnet man als *Hysteresekurve.* Aus Abb. 8.7.3-2 sind die wichtigsten Eigenschaften einer Hysteresekurve ersichtlich: Der Magnetisierungsvorgang ist völlig symmetrisch und hat Sättigungscharakter.

Ein Ferromagnetikum verbleibt nach Abschal-

ten des äußeren Magnetfelds im Zustand eines makroskopischen Permanentmagneten zurück. Seine Magnetisierung in diesem Zustand wird als M_R oder Remanenz bezeichnet.
Erst mit Hilfe einer relativ großen, der Magnetisierung entgegengerichteten magnetischen Feldstärke H_C gelingt es, den Festkörper wieder nahezu vollständig zu entmagnetisieren.
Die eingeschlossene Fläche der Hysteresekurve ist ein Maß für die Energiemenge, die bei einem geschlossenen Ummagnetisierungsvorgang verbraucht wird und als Wärmeenergie freigesetzt wird.
Die für das Ferromagnetikum charakteristischen Werte M_R und H_C werden immer der sogenannten *Grenzhysteresekurve* entnommen. Das ist eine Kurve, die sich bei weiterer Erhöhung der maximal einwirkenden magnetischen Feldstärke nicht weiter verändert. In der Technik verwendet man meist eine Darstellung $B(H)$. Da nun aber für Ferroelektrika $\mu \approx \varkappa$ und $B \approx M$ gilt, sind Angaben, die aus den verschiedenen Darstellungen stammen, nahezu identisch. Analog wie bei den Ferroelektrika steigt die Sättigungsmagnetisierung bei weiterer Erhöhung der Feldstärke noch langsam an. Anhand ihrer Hysteresekurven kann man die Magnetwerkstoffe in zwei unterschiedliche Kategorien einteilen. Dabei reicht es für die Praxis aus, für die *weichmagnetischen Stoffe* eine Koerzitivfeldstärke von $H_C \leqq 2\ \text{Acm}^{-1}$ und für die *hartmagnetischen Stoffe* ein $H_C \geqq 500\ \text{Acm}^{-1}$ zu fordern. Da die besten weichmagnetischen Materialien H_C-Werte von 5 mAcm^{-1} besitzen, ist deren Hysteresefläche außerordentlich schmal. Das heißt aber, daß der Ummagnetisierungsvorgang nur wenig Verlustenergie erfordert. Dagegen ist es bei der Entwicklung von Permanentmagneten besonders wichtig, daß hartmagnetische Werkstoffe eine sehr hohe Koerzitivfeldstärke aufweisen. Die besten Materialien für Dauermagnete erreichen H_C-Werte von 20 kAcm^{-1}. Das besondere Interesse der Magnetforschung richtet sich heute auf intermetallische Verbindungen der seltenen Erden, die H_C-Spitzenwerte und Maximalwerte der Hysteresefläche $\approx 4\, M_R H_C$ besitzen. Auf diese Weise kann der Ummagnetisierungsvorgang so erschwert werden, daß eine magnetische Feldenergie von mehr als $1{,}6 \cdot 10^5\ \text{Wsm}^{-3}$ dafür erforderlich ist. Trotz des hohen Preises dieser Legierungen zeichnen sich hier wesentliche Anwendungen als Permanentmagnete ab. Insbesondere lassen sich mit diesen Materialien die noch kostspieligeren Legierungen Placo und Oerstit ersetzen, die als Grundbestandteile Platin und Cobalt enthalten.

8.7.3.3. Übersicht über die ferromagnetischen Stoffe. In Tab. 8.7.3-3 sind für einige Magnetwerkstoffe die Werte der Sättigungsmagnetisierung und der Curietemperatur zusammengestellt. Da das bei höheren Temperaturen vorliegende Paramagnetikum erst bei der Curietemperatur T_C aus einem spontan magnetisierten Ferromagnetikum entsteht, gilt das Curiegesetz (↑ 8.7.2.2.) in der als *Curie-Weiß-Gesetz* benannten Form

$$(12)\quad \varkappa = C/(T - T_C).$$

Auch hier wird durch diesen Suszeptibilitätsverlauf nur das Paramagnetikum beschrieben, dessen Existenzgebiet sich bei $T > T_C$ erstreckt.
Von den ferromagnetischen Elementen haben nur die Metalle Eisen, Cobalt und Nickel größere praktische Bedeutung erlangt. Als technische Magnetwerkstoffe werden hauptsächlich Legierungen der Eisenmetalle untereinander oder mit anderen nichtferromagnetischen Partnern eingesetzt. Ferromagnetisches Verhalten zeigen aber auch verschiedene Stoffe, die keine ferromangetischen Elemente enthalten. Als Beispiele seien einige intermetallische Verbindungen wie MnAs, MnBi, MnSb oder die Heuslerschen Legierungen Cu_2AlMn und Cu_2SnMn genannt. Für *Dauermagnete* benutzt man hartmagnetische Werkstoffe mit großer Remanenz und Koerzitivfeldstärke, z. B. Chrom- oder Wolframstähle und spezielle Legierungen aus Fe, Ni, Co, Al und Cu. Für Polschuhe und Elektromagneten, zum Abschirmen magnetischer Felder, als Spulen- und Trafokerne bevorzugt man dagegen weichmagnetische Stoffe mit möglichst hoher Anfangs- und auch Maximalsuszeptibilität sowie hoher Sättigungsmagnetisierung. Um unnötige Energieverluste bei der nieder- aber auch oft hochfrequenten Ummagnetisierung zu vermeiden, wird für diese Werkstoffe eine extrem kleine Hysteresefläche gefordert. Viel benutzt werden Reineisen (Armco) sowie Eisen-Nickel- und Eisen-Cobalt-Legierungen mit Zu-

Tab. 8.7.3-3 Sättigungsmagnetisierung M_s einiger Ferromagnetika in Vsm^{-2} bei Zimmertemperatur und Curietemperaturen T_C in K * bei 80 K gemessen

Ferro-magnetikum	Fe	Co	Ni	Gd	Dy	MnBi	Cu_2MnAl	Ni-Zn-Ferrit
M_s	2,15	1,76	0,609	1,37	2,30*	0,76	0,54	0,3
T_C	1043	1400	631	289	105	630	603	500

sätzen von Cu, Cr oder Mo (Muniperm, Mu-Metall, Permalloy). Weitere Energieverluste in Magnetwerkstoffen werden durch in diesen Leitern entstehende *Wirbelströme* hervorgerufen. Man verringert sie durch Unterteilen der Kerne in dünne Bleche, zwischen die isolierende Lagen eingeschoben werden, sowie bei HF-Kernen durch Verpressen isolierter Pulverteilchen. Als Werkstoff für Trafobleche wird hauptsächlich Siliciumeisen mit rund 4% Silicium verwendet. Durch den Siliciumzusatz wird die Suszeptibilität erhöht, während die Wirbelstromverluste wegen der viermal kleineren Leitfähigkeit gegenüber reinem Eisen wesentlich abnehmen. Eine besondere Bedeutung erlangen zunehmend durch Klatschkühlung hergestellte amorphe Mettallegierungen.

8.7.3.4. Spontane Magnetisierung. Die experimentell gesicherten Tatsachen, daß sich ferromagnetische Stoffe in Lösungen und in vielen Verbindungen sowie oberhalb der Curietemperatur wie gewöhnliche Paramagnetika verhalten, während umgekehrt nichtferromagnetische Elemente in bestimmten Legierungen und Verbindungen Ferromagnetismus zeigen, weisen darauf hin, daß der Ferromagnetismus eine Eigenschaft des Festkörpers ist. Die theoretische Deutung geht davon aus, daß die atomaren magnetischen Momente mindestens in räumlichen Bezirken der Größenordnung einige 10 μm^3 einheitlich ausgerichtet sind. Durch vielfältige Untersuchungen wurde sichergestellt, daß der Ferromagnetismus durch die einheitliche Orientierung der Spinmomente von Hüllelektronen nicht abgeschlossener innerer Schalen hervorgerufen wird. Beim Eisenatom sind z. B. die zehn verfügbaren Plätze der 3d-Unterschale mit 6 Elektronen besetzt, von denen 5 eine gleiche Spinorientierung haben, während durch ein 6. Elektron eines der entarteten Energieniveaus doppelt besetzt wird. Da das in Befolgung des Pauliprinzips nur mit Spinabsättigung geschehen kann, trägt ein Eisenatom ein Gesamtmoment von 4 m_B. Von einigen für die Beschreibung des Wesens des Ferromagnetismus geeigneten Theorien soll die *Bändertheorie des Ferromagnetismus* im Abriß dargestellt werden. Im Gitter des metallischen Eisens treten die Elektronen miteinander und mit der Gesamtheit der positiven Atomrümpfe wie in jedem Metall in Wechselwirkung. Es bildet sich eine bestimmte Zustandsdichteverteilung der Elektronenniveaus heraus, die sich in Valenz- und Leitungsbänder einteilen läßt. Die beiden obersten Bänder des Eisens gehen aus den 4*s*- und 3*d*-Zuständen der Eisenatome hervor und überlappen sich. Es kommt auf diese Weise zu einer solchen Besetzung des 3*d*-Bands, daß noch etwa 2,5 Elektronenplätze pro Eisenatom im Band unbesetzt bleiben.

Bei der Behandlung der Reihenfolge in der Besetzung von erlaubten Elektronenniveaus im Rahmen des Bändermodells verfährt man gewöhnlich so, daß man jedes Niveau von niedrigen Energiewerten aus beginnend mit zwei Elektronen entgegengesetzter Spinorientierung auffüllt (↑ 8.2.3.). Für einen paramagnetischen Stoff muß man unter der Wirkung eines Magnetfelds die beiden möglichen Spinorientierungen energetisch unterscheiden. Wie das Abb. 8.7.3-4

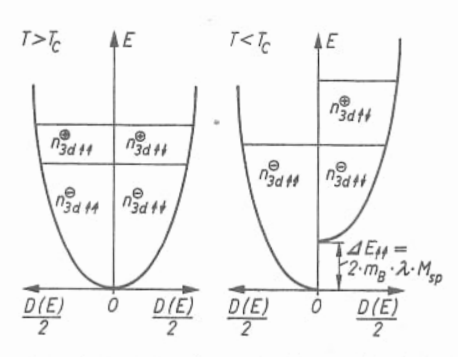


Abb. 8.7.3-4 Besetzung der Energiebänder bei ferromagnetischen Metallen

zeigt, sind in einem Paramagnetikum, also bei $T > T_C$, die beiden Teilbänder $3d(\uparrow\uparrow)$ und $3d(\downarrow\uparrow)$ in gleicher Weise besetzt. Ganz anders sind die Verhältnisse, wenn durch Austauschwechselwirkung zwischen den Elektronen eine *spontane Magnetisierung* zustande kommt. Kennzeichnend für die Stärke dieser spezifisch quantenmechanischen Wechselwirkung, die auch bei der kovalenten Bindung eine entscheidende Rolle spielt, sind Betrag und Vorzeichen des Austauschintegrals, das die Eigenfunktionen und die Abstände der beteiligten Elektronen und Atomrümpfe enthält. Es konnte gezeigt werden, daß nur für positive Werte dieses Integrals die parallele Einstellung vieler benachbarter Elektronenspins energetisch günstig ist, und zwar günstiger als die paarweise antiparallele Einstellung der Elektronenspins. Gerade für die Elemente Fe, Co, Ni sowie für einige seltene Erdmetalle wird das Austauschintegral positiv. In jedem Fall wirkt aber die Wärmebewegung der spontanen Magnetisierung entgegen, so daß in Abhängigkeit von den real vorliegenden strukturellen Bedingungen im Gitter diese bei einem bestimmten T_C-Wert verlorengeht. Bei den seltenen Erdmetallen ist die Austauschwechselwirkung so schwach, daß nur bei genügend tiefen Temperaturen ferromagnetisches Verhalten in einem Gitter oder Untergitter auftritt.

Man kann nun einfach die spontane Magnetisierung M_{sp} in einem bestimmten Volumenelement als gegeben ansetzen und davon ausgehen,

daß am Ort jedes magnetischen Dipols das lokale Magnetfeld λM_{sp} herrscht. Unter dieser Voraussetzung findet man eine Absenkung der Energieniveaus aller Elektronen mit Parallelorientierung $\Delta E_{\uparrow\uparrow} = 2m_B\lambda M_{sp}$. Die zwei Teilbänder werden also (↑ Abb. 8.7.3-4) asymmetrisch bis zum Ferminiveau aufgefüllt. Als Ergebnis entsteht ein makroskopisch nachweisbarer Magnetisierungszustand des Festkörpers.

8.7.3.5. Ferromagnetische Domänen. Die Vorstellung von der spontanen Magnetisierung widerspricht scheinbar dem ferromagnetischen Verhalten von frisch aus der Schmelze hergestellten Magnetwerkstoffen oder dem Verhalten weichmagnetischer Stoffe, die kein magnetisches, nach außen hin wirksames Gesamtmoment haben. Wie WEISS zuerst erkannte, ist die spontane Magnetisierung auf Volumenbereiche der Größenordnung einige 10 μm^3 begrenzt. Innerhalb eines Ferromagnetikums sind viele solcher *Weißschen Bezirke* oder *ferromagnetischer Domänen* verschiedener Orientierung vorhanden (Tafel 47). Die Weißschen Bezirke sind nicht mit den Kristallkörpern polykristalliner Materialien zu verwechseln. Sie treten vielmehr auch in Einkristallen auf, und bei polykristallinen Ferromagnetika umfaßt das Korn meist mehrere Domänen. Die Einteilung eines Festkörpers in Domänen hat energetische Ursachen. Es läßt sich z. B. zeigen, daß für Eisen die optimale Domänengröße $(2\ \mu m)^3$ beträgt. Dieser Wert kann berechnet werden, wenn die konkurrierenden Einflüsse der magnetischen Feldenergie im Festkörper, im Außenraum und an den Domänenwänden in Rechnung gestellt werden. Die *Domänenwände* oder *Blochwände* sind Grenzflächen innerhalb eines Ferromagnetismus, die verschiedene Magnetisierungsrichtungen voneinander trennen. Wie bei den Ferroelektrika erfolgt der Übergang von einer Orientierung in die andere stetig. Eine Blochwand umfaßt deshalb etwa einen Bereich von 100 bis 1000 Netzebenen. In Abb. 8.7.3-5

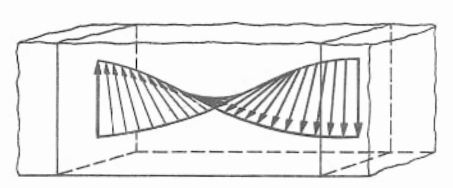


Abb. 8.7.3-5 Blochwand

ist dargestellt, wie man sich die Lage der magnetischen Dipole in einer Blochwand vorzustellen hat. An dieser Grenzfläche tritt als Folge des Wechsels der Orientierung der Dipollagen eine hohe Feldstärke auf. An der geschliffenen und geätzten Oberfläche eines Ferromagnetikums lassen sich daher Domänenwände sichtbar machen. Das geschieht durch Ablagerung feinsten Eisenpulvers aus einer Suspension.

8.7.3.6. Magnetisierung eines Dauermagneten. Bei der Magnetisierung eines Ferromagnetikums in einem äußeren Magnetfeld wachsen durch Verschiebung der Domänenwände im Gitter zunächst die Domänen auf Kosten der anderen, die von vornherein bezüglich der Richtung des äußeren Felds eine günstige Orientierung aufweisen. In einem ideal gebauten Gitter ist die Verschiebung einer Domänenwand leicht möglich. Anders sieht das aus, wenn der Festkörper eine bestimmte Menge von Defekten enthält. Unter realen Bedingungen hakt sich die betreffende Blochwand mehr oder weniger schnell in Gitterdefekten fest. Ihre weitere Verschiebung ist prinzipiell möglich, aber es wird eine wesentlich höhere Energie für diesen Vorgang verbraucht. Daraus erklärt sich der Unterschied zwischen hartmagnetischen und weichmagnetischen Stoffen und für die irreversiblen Erscheinungen der ferromagnetischen Hysterese. Auf dem irreversiblen Anteil der Verschiebung von Blochwänden beruht die Remanenz. Man kann sie wesentlich erhöhen, wenn durch Fremdstoffe Gitterstörungen erzeugt werden, die die Drehung der Elementarmagnete behindern. In den Bereich der Sättigungsmagnetisierung wird ein Ferromagnetikum gebracht, wenn das angelegte Magnetfeld eine Stärke erreicht, daß ganze Domänen in ungefähr die richtige Orientierung gebracht werden können. Ein solcher Orientierungsvorgang erfordert verständlicherweise viel höhere Kräfte als die Drehung eines Einzeldipols. Auf dieser Basis lassen sich deshalb Dauermagneten hoher Remanenz und hoher Koerzitivfeldstärke herstellen, indem Sinterkörper (Maniperm) aus sehr fein gemahlenen ferromagnetischen Stoffen verwendet werden. Die verpreßten ferromagnetischen Teilchen umfassen dann nur noch je eine Domäne, die nicht über die Korngrenze hinaus ausgedehnt werden kann. Der Magnetisierungvorgang solcher *Sintermagnete* beruht dann nahezu ausschließlich auf der Änderung der Orientierung ganzer Domänen.

Das Wachstum der Weißschen Bezirke und auch die Drehung ihrer Magnetisierungsrichtung erfolgen bei Erhöhung eines äußeren Magnetfelds nicht kontinuierlich, sondern in kleinen Sprüngen. Diese kleinen Unregelmäßigkeiten sind auch in der Kurve $M(H)$ sichtbar. Bringt man z. B. einen Eisenstab während der Magnetisierung in eine Induktionsspule, dann werden in dieser kleine Spannungsstöße induziert. Nach entsprechender Verstärkung kann man sie in einem Lautsprecher als prasselndes Geräusch wahrnehmen. Nach ihrem Entdecker BARKHAUSEN nennt man diese Erscheinung *Barkhauseneffekt*.

8.7.3.7. Spinwellenmagnonen. Eine streng par-

allele Orientierung magnetischer Momente in einem Ferromagnetikum ist nur am absoluten Nullpunkt gewährleistet. Bei höheren Temperaturen, die noch weit unterhalb der Curietemperatur liegen können, führen die magnetischen Dipole um die Magnetisierungsrichtung in einer Domäne Schwingungen aus. Da auch hier die Erhaltung des Drehimpulses einen wesentlichen Einfluß ausübt, sind solche Schwingungen immer Drehschwingungen des magnetischen Moments um die Magnetisierungsrichtung. Im Gitter des Festkörpers sind die Schwingungszustände miteinander gekoppelt, so daß zwischen benachbarten Gitterplätzen bestimmte Phasenbeziehungen herrschen (Abb. 8.7.3-6). zusätzlichen Energieaufwand, der als *Anisotropieenergie* bezeichnet wird. Bei einer Magnetisierung in energetisch ungünstiger Richtung ordnen sich die Gitterbausteine etwas um. Die Folge sind kleine Längenänderungen des Festkörpers. Man kann sie z. B. bei polykristallinen Eisen- oder Nickelproben beobachten. Beim Nickel tritt gegenüber der ursprünglichen Länge bei Sättigungsmagnetisierung in der entsprechenden Richtung eine Verkürzung um 0,0035% ein. Magnetisiert man einen Nickelstab in einem hochfrequenten Wechselfeld, so ergeben sich Längsschwingungen, die sich aufschaukeln, wenn es zur Resonanz zwischen der Frequenz der mechanischen Eigenschwingungen und der Frequenz des Wechselfelds kommt. Dieser Effekt wird als *Magnetostriktion* bezeichnet. Auf dieser Grundlage ist der Bau *magnetostriktiver Schwingungsgeber*, z. B. in Form von robusten Ultraschallquellen sowie von Lötkolben für das Löten von Aluminium möglich. Für weichmagnetische Werkstoffe ist der *Magnetostriktionseffekt* von großem Interesse, weil der dynamische Betrieb in magnetischen Werkstoffen enorme mechanische Spannungen erzeugen kann, deren Nichtbeachtung zu Schwingungen und Zerstörungen von Transformatoren, Motoren und Generatoren führen würde.

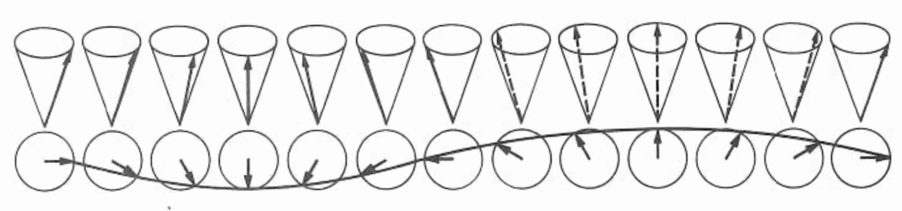


Abb. 8.7.3-6 Spinwelle

Infolgedessen läßt sich die Summe der Bewegungen des gesamten Spinsystems als Überlagerung thermisch angeregter *Spinwellen* beschreiben, die die Domäne nach allen Richtungen durchlaufen. Die Verhältnisse ähneln in bestimmter Weise denen von Gitterschwingungen, die als gequantelte Quasiteilchen aufgefaßt und als Phononen bezeichnet werden. Auch die Spinwellen unterliegen bestimmten Quantenvorschriften. Man verwendet daher gleichfalls mit Erfolg die Konzeption der Quasiteilchen, die in diesem Fall als *Magnonen* bezeichnet werden. Magnonen unterliegen der Bose-Einstein-Statistik. Sie sind die niedrigsten angeregten Zustände über dem Grundzustand eines Ferro-, Antiferro- oder Ferrimagnetikums. Ihr Nachweis erfolgt mit Hilfe der inelastischen Neutronenstreuung, wobei die magnetische Dipol-Dipol-Wechselwirkung zur Anregung von Spinwellen im Gitter führt.

8.7.3.8. Magnetische Anisotropie – Magnetostriktion. Innerhalb ferromagnetischer Kristalle gibt es Richtungen, in denen sich die spontane Magnetisierung bevorzugt ausbildet. Meist handelt es sich bei diesen Richtungen um niedrigindizierte Richtungen, wie z. B. beim Eisen um die [100]-, [110]- oder [111]-Richtung. In einer bestimmten Richtung ist dann immer der Magnetisierungsvorgang mit geringstem Energieaufwand möglich (Abb. 8.7.3-7). Bei Eisen ist das die [100]-Richtung, beim Nickel dagegen die [111]-Orientierung. Jedes Herausdrehen der Magnetisierung aus diesen bevorzugten Richtungen durch ein äußeres Feld erfordert einen

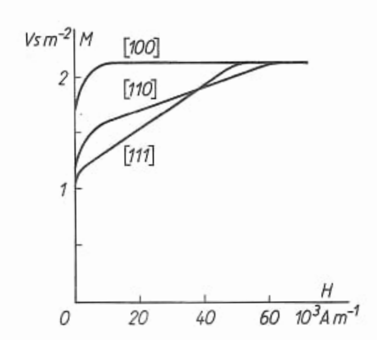


Abb. 8.7.3-7 Magnetisierung des Eisens in Abhängigkeit von der äußeren Feldstärke *H* bezüglich verschiedener Kristallorientierungen

8.7.4. Antiferro- und Ferrimagnetismus

8.7.4.1. Antiferromagnetismus. In Antiferromagnetika besteht zwar eine spontane Ordnung der atomaren Momente, diese sind aber im Gegensatz zu einem Ferromagnetikum in benachbarten Gitterebenen antiparallel orientiert. Deshalb verhält sich ein antiferromagnetischer Festkörper in erster Linie makroskopisch wie ein Paramagnetikum. Aus diesem Grunde wurde die Erscheinung auch erst 1932 von NÉEL entdeckt. Unterhalb einer bestimmten Temperatur, der Néeltemperatur, ist es für verschiedene Festkörper energetisch günstiger, in einen Zustand antiparalleler Spinorientierung überzugehen. Zu den antiferromagnetischen Stoffen gehören z. B. MnO, MnS, MnF_2, FeO, $FeCl_2$, CoO, Cr, Mn, Nd, Sm und einige seltene Erden. Kennzeichnend für diese Stoffe, deren Néeltemperaturen oft weit unterhalb der Zimmertemperatur liegen, ist ein Suszeptibilitätsverlauf in Abhängigkeit von der Temperatur mit einem deutlichen Maximum (Abb. 8.7.4-1). Der Anstieg der Suszeptibilität ist mit einer Lockerung der spontanen Ordnung durch Wärmeschwingungen zu erklären. Das Maximum liegt in der Umgebung der Temperatur, wo der Ordnungszustand zusammenbricht. Oberhalb der Néeltemperatur sind *Antiferromagnetika* dann paramagnetisch. Der Verlauf $\varkappa(T)$ befolgt oberhalb der Umwandlungstemperatur ein Curie-Weiß-Gesetz. Es gilt $\varkappa = C/(T + \Theta)$.

Die Temperatur Θ ist eine charakteristische Temperatur, die nicht mit der Néeltemperatur identisch ist. Sie muß dem Kurvenverlauf $\varkappa(T)$ entnommen werden. In dem gründlich untersuchten Beispiel des Magnesiumoxids liegen besonders einfache Verhältnisse vor, da die Mn-Ionen mit ihren Momenten innerhalb einer Gitterebene die gleiche Richtung aufweisen, während sie in den unmittelbar benachbarten Gitterebenen entgegengesetzt gerichtet sind (Abb. 8.7.4-2). Bei komplizierter gebauten Antiferromagnetika können auch mehr als zwei Untergitter am Zustandekommen einer wechselseitigen Kompensation lokaler spontaner Magnetisierungsvorgänge beteiligt sein.

8.7.4.2. Ferrimagnetismus. Ferrimagnetismus liegt dann vor, wenn keine vollständige Kompensation der Wirkungen magnetischer Untergitter in einem Festkörper vorhanden ist. Das kann auf verschiedene Art realisiert sein, z. B. durch die verschiedene Anzahl der magnetischen Dipole oder durch ihre verschiedene Größe in den Untergittern. *Ferrimagnetika* besitzen deshalb eine resultierende spontane Magnetisierung. Sie verhalten sich makroskopisch ähnlich wie Ferromagnetika.

Ferrimagnetika haben als *Ferrite* eine große praktische Bedeutung erlangt. Ferrite mit Spinellgitter besitzen die allgemeine Zusammensetzung $Me(II)Fe_2(III)O_4$. Das Symbol Me(II) bedeutet dabei ein zweiwertiges Metallion wie Fe^{2+}, Mn^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+} oder Mg^{2+}. Es lassen sich auch *Mischferrite* der Zusammensetzung $Me_x^{(1)}(II)\,Me_{1-x}^{(2)}(II)\,Fe_2(III)O_4$ oder $Me_x^{(1)}(II)Me_y^{(2)}(II)Me_{1-x-y}^{(3)}(II)Fe_2(III)O_4$ durch Sintern aus Oxidpulvern bei 1200 °C herstellen. Ferrite haben den Vorteil eines hohen

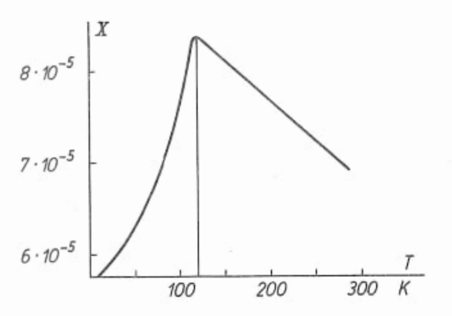


Abb. 8.7.4-1 Temperaturverlauf der magnetischen Suszeptibilität $\varkappa$ von MnO

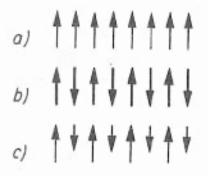


Abb. 8.7.4-2 Spinorientierung in benachbarten Gitterebenen; a) Ferromagnetikum, b) Antiferromagnetikum, c) Ferrimagnetikum

spezifischen Widerstands, so daß Wirbelstromverluste weitgehend ausgeschlossen sind. Für Speicherkerne kommen vor allem Ni-Zn-Ferrite und Mg-Mn-Ferrite in Betracht. Für die Mikrowellentechnik wurden speziell die *ferrimagnetischen Granate* entwickelt. Granate haben die allgemeine Summenformel $A_3B_2C_3 \cdot O_{12}$. Bekannteste Vertreter sind z. B. Yttriumgranat ($Y_3Fe_2Fe_3O_{12}$, Abkürzung YIG) und Gadoliniumgranat ($Gd_3Fe_2Fe_3O_{12}$).

Es soll das Zustandekommen des Ferrimagnetismus für den relativ übersichtlichen Fall des Nickelferrits $NiFe_2O_4$ näher erläutert werden. Dieser Werkstoff kristallisiert wie die übrigen *Ferrospinelle* im Spinellgitter, dessen Elementarzelle acht Formeleinheiten (32 O^{2-}-, 16 Fe^{2+}- und 8 $Me(II)^{2+}$) umfaßt (Abb. 8.7.4-3). Die Metallionen sind auf zwei verschiedenen Typen von Gitterplätzen untergebracht. In den acht Tetraederplätzen sind die Metallionen tetraedrisch von je vier Sauerstoffionen umgeben, in den sechzehn Oktaederplätzen dagegen von je sechs Sauerstoffionen. Beim Nickelferrit befinden

sich 8 Fe^{2+}-Ionen auf den Tetraeder- und die anderen 8 Fe^{2+} zusammen mit den 8 Ni^{2+} auf den Oktaederplätzen. Die Fe^{2+}-Ionen haben ein magnetisches Moment von 5 m_B, die Ni-Ionen von 2 m_B, und die Sauerstoffionen besitzen

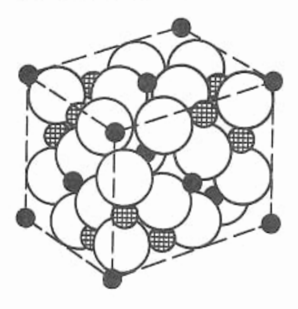


Abb. 8.7.4-3 Elementarzelle des Spinellgitters; *weiß* Sauerstoffionen, *schwarz* Metallionen auf Tetraederplätzen (A-Plätze), *schraffiert* Metallionen auf Oktaederplätzen (B-Plätze)

kein magnetisches Moment. Zwischen den Oktaeder- und Tetraederplätzen besteht eine starke antiferromagnetische Kopplung, so daß sich die magnetischen Momente auf den entsprechenden Plätzen zueinander antiparallel orientieren. Einem Tetraederplatz mit 5 m_B stehen demnach je zwei Oktaederplätze mit $(5 + 2)\, m_B = 7\, m_B$ gegenüber. Als Differenz ergibt sich ein Ferrimagnetismus von 2 m_B pro Formeleinheit oder 16 m_B pro Elementarzelle.

8.7.5. Magnetspeicher

8.7.5.1. Flächenspeicher. Die Remanenz eines Magnetwerkstoffs in zwei Zuständen mit entgegengesetzter Magnetisierungsrichtung kann zur Speicherung der binären Zahlen O und L und damit von 1 bit Information dienen. Die Information wird ohne Leistungsaufwand festgehalten. Die elektronische Datenverarbeitung erfordert Magnetspeicher großer Kapazität, für einen Rechner mittlerer Größe z. B. von 100 kbit bis 100 Mbit. Dabei muß die Konstruktion des Speichers kleine Zugriffszeiten beim Auffinden der Information zulassen.

Flächenspeicher werden als externe Massenspeicher sowie als *Arbeitsspeicher* bei mäßigen Forderungen an die Zugriffszeit eingesetzt. Der Speichereffekt beruht wie bei Tonbändern auf der Magnetisierung kleiner Gebiete einer hartmagnetischen Schicht. Der hartmagnetische Werkstoff wird polykristallin auf einem Band, einer Trommel oder einer Platte elektrolytisch oder durch Aufdampfen abgeschieden. Das Einschreiben der Information erfolgt über einen von der elektronischen Anlage gesteuerten Aufnahmekopf. Im Lesekopf erzeugen die verschieden magnetisierten Gebiete beim Vorbeigleiten positive bzw. negative Spannungsimpulse, die entsprechend weiterverarbeitet werden. Mit sehr schnell rotierenden *Trommelspeichern* lassen sich Zugriffszeiten von 5 bis 500 ms erreichen.

8.7.5.2. Ferritkernspeicher. Ein *Ferritkernspeicher* ist aus ringförmigen Ferritkernen von 0,2 bis 0,8 mm Durchmesser aufgebaut, die von einem System sich kreuzender Drähte durchzogen sind (Abb. 8.7.5-1). Es sind die wichtigsten Arbeitsspeicher, Durch jeden Kern verlaufen je zwei Einschreibleitungen und die Leseleitung. Das gesamte System der Kerne in einer Ebene wird als *Speichermatrix* oder *Speicherebene* bezeichnet (Tafel 78). Eine Speichereinheit besteht aus vielen Speicherebenen. Um beispielsweise in den Kern mit den Koordinaten y_3 und x_2 eine Information einzuschreiben, werden von der Ansteuerelektronik an die y_3-Leitung und die x_2-Leitung Stromimpulse gegeben, die so bemessen sind, daß nur die Summe der von ihnen erzeugten Magnetfelder ausreicht, in der gewünschten Richtung die Koezitivfeldstärke zu übertreffen. Eine Ummagnetisierung erfolgt also nur dort, wo sich beide Leitungen kreuzen. Für eine einwandfreie Funktion der Kerne muß ihre Hysteresekurve nahezu die Gestalt eines Rechtecks haben. Zum Lesen der gespeicherten Information fragt man die Ferritkerne der Reihe nach mit Impulsfolgen ab. Eine Ummagnetisierung wird dabei in der Leseleitung als Spannungsimpuls wirksam. Allerdings wird dabei die gespeicherte Information gelöscht, so daß sie in Verbindung mit dem Lesevorgang sofort wieder einzuschreiben ist. Die Arbeitsgeschwindigkeit des Ferritkernspeichers wird von der *Zykluszeit* bestimmt. Das ist die Zeitspanne zwischen dem Abfragen und dem Wiedereinschreiben der Information. Sie ist hauptsächlich durch die endliche Dauer des Ummagnetisierungsprozesses gegeben und sinkt mit abnehmendem Durchmesser der Kerne.

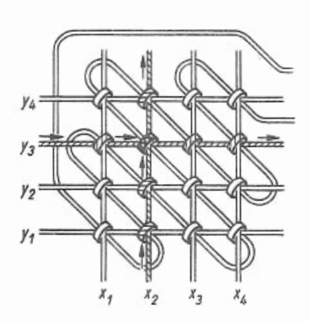


Abb. 8.7.5-1 Teil der Speichermatrix eines Ferritkernspeichers (s. a. Tafel 78)

8.7.5.3. Magnetische Dünnschichtspeicher. Aussichtsreich erscheint auch die Anwendung magnetischer Dünnschichtspeicher. Sie lassen sich in den zwei Varianten Magnetblasenspeicher und DOT-Speicher herstellen. Beiden Typen liegt zugrunde, daß in einer dünnen Schicht eines Magnetwerkstoffs eine geordnete Folge von Domänen entsteht, die als Elemente eines Speichers dienen können. Beim *Magnetblasen-* oder *Bubblespeicher* wird der Magnetwerkstoff so abgeschieden, daß seine bevorzugte Richtung für spontane Magnetisierung senkrecht auf dem Substrat steht. In geeigneten Materialien können zylinderförmige Domänen erzeugt werden, die sich längs einer vorgegebenen Morphologie des Dünnschichtelements, z. B. durch Ätzen erzeugt, bewegen können. Damit sind Speicher in der Form eines Schieberegisters möglich. Wesentliches Prinzip beider Speichertypen ist, daß die Informationsspeicherung in einer Schicht erfolgt, die in einer Richtung bis zur Sättigung magnetisiert ist. Am Ort der Aufzeichnung wird eine Umkehr der Magnetisierungsrichtung erzwungen. Das geschieht wieder durch gekreuzte Leiterbahnen, die durch eine isolierende Zwischenschicht voneinander isoliert sind. Allein durch die Wirkung äußerer Magnetfelder können die Domänen innerhalb des Speichermediums bis zu einem Lesekopf, z. B. einer Induktionsspule verschoben werden. Beim *DOT-Speicher* [engl. domain tip propagation storage device] liegt die Vorzugsrichtung der Magnetisierung in der Schichtebene. Ein DOT-Speicher wird auf einer Glasplatte von einigen Quadratzentimetern Fläche abgeschieden. Zuerst wird eine mäanderförmige Aluminiumbahn aufgetragen, dann eine 160 nm dicke Schicht einer Ni/Fe/Co-Legierung und darauf scheidet man Leiterschleifen zum Einschreiben ab. Es werden Packungsdichten von 1 bis 3 kbit cm^{-2}, Einschreibgeschwindigkeiten von 0 bis 10 Mbit s^{-1} und Zugriffszeiten im Bereich von Millisekunden erreicht. Die Domänen lassen sich mit 10 ms^{-1} durch das Speichermedium hindurch verschieben. Bei den zylindrischen Dünnschichtspeichern oder *magnetischen Drahtspeichern* scheidet man die ferromagnetische Schicht elektrolytisch auf dünnen Kupferdrähten ab und teilt sie in einem Ätzprozeß in zahlreiche ringförmige Gebiete ein. Die fertigen Drähte werden in einem automatisierten Webvorgang zusammen mit einem System von Paralleldrähten zur Speichermatrix verarbeitet.

8.8. Optische Eigenschaften der Festkörper

8.8.1. Reflexion, Brechung und Absorption

8.8.1.1. Einführung. Charakteristische optische Eigenschaften fester Stoffe, beispielsweise der Glanz und die typische Färbung von Edelsteinen, die spiegelnde Lichtreflexion von Metallen oder die Durchsichtigkeit von Gläsern, haben frühzeitig das Interesse der Menschen gefunden. Auch Anwendungen dieser Eigenschaften, etwa zur Herstellung von Spiegeln oder Fenstern, reichen bis in die Anfänge der Zivilisation zurück. Mit dem Aufstreben der Naturwissenschaften lernte man es, die wesentlichen Erscheinungen beim Durchgang von Licht durch Festkörper quantitativ zu beschreiben. Man benutzt dazu Materialgrößen wie Brechzahl und Absorptionskoeffizient, die für unterschiedliche Stoffe verschiedene Werte besitzen und Funktionen der Wellenlänge bzw. der Wellenzahl des Lichts sind. Diese phänomenologische Beschreibung ermöglicht es, das optische Verhalten auch komplizierter Systeme zu berechnen, und sie bildet daher die Grundlage für die Konzipierung von optischen Geräten. Erst auf dem Boden der modernen Festkörperphysik war es möglich, die optischen Materialgrößen auch atomistisch zu interpretieren. Dadurch eröffneten sich neue Möglichkeiten zur gezielten Beeinflussung des optischen Verhaltens der Werkstoffe, und darüber hinaus wurde der Weg für neue wichtige Anwendungen bereitet, die auf einer Nutzung der spezifischen Wechselwirkung zwischen Lichtwellen und dem Atomverband fester Stoffe aufbauen.

Prinzipiell wichtig war die Erkenntnis, daß das sichtbare Licht einen Ausschnitt aus dem Spektrum der *elektromagnetischen Wellen* (↑ 2.2.4.) verkörpert. Infolgedessen betrachtet man in der Festkörperoptik nicht nur die Phänomene im sichtbaren Spektralgebiet, sondern bezieht die physikalisch im Prinzip gleichartigen Vorgänge in anderen Wellenlängenbereichen, besonders den angrenzenden Gebieten des Infraroten und des Ultravioletten, ein. Zugleich liefert die elektromagnetische Theorie den Schlüssel für das Verständnis des Mechanismus der Wechselwirkung zwischen Lichtwellen und Festkörpern. Danach löst die zeitlich und räumlich periodisch veränderliche elektrische Feldstärke $\boldsymbol{E} = \boldsymbol{E}_0 \exp i(\boldsymbol{kr} - \omega t)$ der Lichtwelle mit der Kreisfrequenz ω und dem Wellenzahlvektor $\boldsymbol{k}$ am Ort $\boldsymbol{r}$ erzwungene Schwingungen der elektrisch polarisierbaren atomaren Bausteine aus (s. a. 8.5.2.). Auf diese Weise wird eine periodisch veränderliche elektrische Polarisation hervorgerufen, die atomistisch als Schwingungsbewegung von Ionen in Form *optischer Phononen* (↑ 8.1.2.2.) oder von Elektronen relativ zu den Atomrümpfen, seltener auch von Dipolen molekularer Bausteine aufzufassen ist. Nach den Gesetzen der Elektrodynamik müssen die schwingenden Ladungen als Oszillatoren selbst Licht ausstrahlen. Aus dem Feld der einfallenden Lichtwelle und den Beiträgen der atomaren Oszillatoren baut sich im Festkörper ein Wellen-

feld auf, das den Durchgang der elektromagnetischen Welle durch das Medium beschreibt. Gegenüber der vorherrschenden Wechselwirkung über das elektrische Feld der Lichtwelle tritt die magnetische Wechselwirkung zurück, so daß sie, außer bei ferromagnetischen Stoffen, vernachlässigt werden kann. Im Falle von Einkristallen ergeben sich Komplikationen, weil die optischen Eigenschaften nichtkubischer Kristalle richtungsabhängig, also anisotrop sind. Bei polykristallinen Materialien mitteln sich die anisotropen Eigenschaften jedoch heraus, und auch die als Werkstoffe der Optik besonders wichtigen Gläser und Plaste verhalten sich wegen des weitgehend regellosen molekularen Aufbaus gewöhnlich isotrop. Im folgenden wird zunächst das Verhalten optisch isotroper Festkörper betrachtet, während die mit der *Anisotropie* zusammenhängenden Erscheinungen später behandelt werden (↑ 8.8.2.).

Natürlich ist bei optischen Vorgängen stets der Welle-Teilchen-Dualismus (s. a. 1.2.4.) des Lichts zu beachten. Der Quantencharakter des Lichts tritt insbesondere dann hervor, wenn Einzelheiten der Absorption oder Emission betrachtet werden. Das ist vor allem bei den *Fotoeffekten* (↑ 8.8.3.) unumgänglich, aber auch die chemischen Wirkungen des Lichts, darunter die *Grundprozesse der Fotografie* (↑ 8.8.4.), beruhen vorwiegend auf Quantenphänomenen.

8.8.1.2. Optische Materialgrößen. Trifft ein Lichtbündel auf die als glatt angenommene Oberfläche eines isotropen Festkörpers auf, so wird ein Teil der Strahlung reflektiert, während der Rest in den Festkörper eindringt und mehr oder weniger stark absorbiert wird. Für die Winkelbeziehungen zwischen den drei Strahlen (Abb. 8.8.1-1) gilt nach dem *Reflexionsgesetz* und dem *Snelliusschen Brechungsgesetz*

(1) $\alpha_0 = \alpha_0', \quad \sin\alpha_0/\sin\alpha_1 = n_1/n_0.$

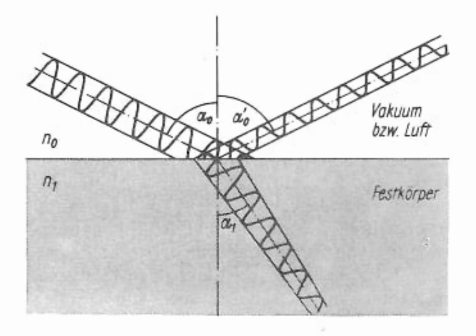


Abb. 8.8.1-1 Verhalten einer Lichtwelle an der Grenzfläche Vakuum/Festkörper; die abnehmende Amplitude im Festkörper veranschaulicht die Absorption

Die Winkel werden in der Optik stets relativ zur Flächennormalen angegeben. Mit n_1 wird durch [8.8.(1)] als wichtigste optische Materialgröße die Brechzahl bzw. der Brechungsindex eingeführt. Für das Vakuum gilt $n_0 = 1$, und näherungsweise kann dieser Wert auch beim Lichteinfall aus Luft übernommen werden. Aus der elektromagnetischen Theorie des Lichts folgt, daß die Brechzahl gleich dem Verhältnis der Vakuumlichtgeschwindigkeit c zur Ausbrei-

Tab. 8.8.1-2 Brechzahlen verschiedenartiger fester Stoffe

Stoff	Brechungsindex	Wellenlänge in nm
Gläser		
Fluor-Kron-Glas	1,44 bis 1,49	587,6
Kron-Glas	1,49 bis 1,54	587,6
Schwer-Flint-Glas	1,62	587,6
Lanthan-Kron-Glas	1,69	587,6
Quarz-Glas	1,458	589,3
transparente Kunststoffe		
Polytetrafluorethylen (PTFE)	1,35 bis 1,37	589,3
Polyethylen	1,51	589,3
Polystyren	1,59 bis 1,60	589,3
Epoxydharze	1,57 bis 1,58	589,3
Polyvinylchlorid	1,54 bis 1,95	589,3
Halbleiter		
CdS	2,506	589,3
Se	2,93	588,0
Si	4,18	589
GaAs	4,04	546,1
Ge	5,3	546,1
SiO	1,43	589,3
ZnS (Zinkblende)	2,366	589,3
Salze		
NaF	1,3258	587,7
	1,3930	186
NaCl	1,5441	589
KCl	1,4899	577,0
	1,7077	202,6
KNO_3 (Kalisalpeter)	1,3346	589,3
	1,367	253,0
NH_4Cl (Salmiak)	1,6422	589
$CaSO_4$ (Anhydrit)	1,569	589,3
$CaCO_3$ (Kalkspat)	1,658	589
$CuSO_4 \cdot 5\,H_2O$ (Kupfervitriol)	1,5141	589
einige Oxide		
Li_2O	1,644	589
BeO (Bromellit)	1,719	589
MgO (Periklas)	1,7366	589
CaO	1,837	589

tungsgeschwindigkeit im Medium v ist, und es ergibt sich darüber hinaus ein einfacher Zusammenhang mit der Dielektrizitätskonstanten ε und der Permeabilität μ des Materials:

(2) $n = c/v = (\varepsilon\mu)^{1/2}$.

Die außer bei Ferromagnetika stets brauchbare Näherung $n \approx \sqrt{\varepsilon}$ ist gerechtfertigt, weil die Permeabilität der meisten Stoffe nur wenig von Eins verschieden ist. Die hier auftretende Dielektrizitätskonstante darf jedoch nicht mit derjenigen gleichgesetzt werden, die bei statischen, nieder- oder hochfrequenten elektrischen Messungen erhalten wird, sondern für die extrem hohen Frequenzen der Lichtwellen gelten meist kleinere Werte. Allgemein hängen n und damit auch ε von der Frequenz des benutzten sichtbaren oder unsichtbaren Lichts ab, ein Verhalten, das man als *Dispersion* bezeichnet.

Unter Verwendung von monochromatischem Licht läßt sich die Brechzahl mit Hilfe geeigneter optischer Meßanordnungen sehr genau bestimmen. Für eine Reihe fester Stoffe sind die bei den angegebenen Lichtwellenlängen gemessenen Werte in Tab. 8.8.1-2 zusammengestellt. Bei Kenntnis der Brechzahlen läßt sich die Lichtausbreitung beim Durchgang der Strahlen durch die Grenzfläche zwischen zwei beliebigen Medien mit den Brechzahlen n_1 und n_2 mit dem verallgemeinerten Brechungsgesetz beschreiben: $\sin\alpha_1/\sin\alpha_2 = n_2/n_1$. Daneben wird die Brechzahl häufig zur Charakterisierung der Reinheit von Stoffen benutzt. Beim Übergang von Licht aus einem optisch dichteren Medium mit größerer Brechzahl in ein Medium mit kleinerer Brechzahl ($n_1 > n_2$), also beispielsweise beim Austritt von Licht aus einem Festkörper in Luft, tritt oberhalb des durch $\sin\alpha_1 = n_2/n_1$ gegebenen Grenzwinkels Totalreflexion ein.

Im allgemeinen sind die stofflichen Medien, darunter die Festkörper, für das Licht nicht ideal durchlässig, sondern es findet eine mehr oder weniger starke Absorption oder auch eine Lichtstreuung statt. Um das Verhalten absorbierender Medien zu beschreiben, führt man als weitere Materialgrößen den Absorptionsindex a ein, der die Dämpfung der Lichtamplitude angibt. Brechzahl und Absorptionsindex werden häufig zu dem komplexen Brechungsindex $n' = n + \mathrm{i}\,a$ zusammengefaßt. Aus der elektromagnetischen Lichttheorie folgt für n' der Ausdruck

(3) $n' = k/k_0 = (\varepsilon + \mathrm{i}\sigma/\omega\varepsilon_0)^{1/2}$,

wobei σ die elektrische Leitfähigkeit bezeichnet. Die Lichtabsorption ist demnach mit der elektrischen Leitfähigkeit im betrachteten Frequenzgebiet verknüpft. Weiter gibt k/k_0 das Verhältnis der Wellenzahl im absorbierenden Medium k zur Vakuumwellenzahl $k_0 = \omega/c$ an. Die Brechzahl und der Absorptionsindex entsprechen dem Real- bzw. dem Imaginärteil von 8.8.(3). Damit läßt sich für die Wellenzahl im Medium $k = k_0 n + \mathrm{i}k_0 a$ schreiben. Setzt man diesen Ausdruck in die eindimensionale Wellengleichung der elektrischen Feldstärke $\boldsymbol{E}$ einer in x-Richtung fortlaufenden Lichtwelle ein, so folgt $\boldsymbol{E} = \boldsymbol{E}_0 \exp -(k_0 a x) \exp \mathrm{i}(k_0 n x - \omega t)$. Als Lösung für das absorbierende Medium ergeben sich also gedämpfte Wellen, und der Absorptionsindex a ist ein Maß für den exponentiellen Dämpfungsfaktor. Zahlenwerte der Brechzahl und des Absorptionsindex sind für einige feste Elemente in Tab. 8.8.1-3 angegeben.

8.8.1.3. Intensitäten der reflektierten und gebrochenen Lichtanteile. Wie FRESNEL bereits zu Beginn des vorigen Jahrhunderts zeigte, lassen sich mit Hilfe der optischen Materialgrößen nicht nur die Brechungswinkel, sondern auch die relativen Intensitäten der reflektierten und gebrochenen Lichtanteile angeben. Allerdings muß man die einfallende Lichtwelle dazu in zwei linear polarisierte Komponenten zerlegen. Bei der mit s indizierten Welle schwingt der elektrische Feldstärkevektor senkrecht zur Einfallsebene, die vom einfallenden Strahl und der Flächennormalen aufgespannt wird, während er bei der anderen mit p gekennzeichneten Komponente parallel zu dieser Ebene schwingt. Bei der modernen Begründung auf der Grundlage der elektromagnetischen Lichttheorie geht man von

Tab. 8.8.1-3 Brechzahl n und Absorptionsindex χ einiger Elemente bei $\lambda = 589$ nm

Stoff	n	χ
Na	0,048	1,86
K	0,07	1,5
Be	2,6	2,27
Ca	0,29	3,31
Mg	0,37	4,42
Al	1,44	5,23
Ta	2,18	2,16
Cr	3,59	4,5
Mo	1,99	1,126
W	3,46	3,25
Mn	2,41	3,88
Fe	2,36	3,20
Co	2,0	4,0
Ni	1,74	3,39
Pt	1,82	3,78
Cu	0,617	2,63
Ag	0,181	3,67
Au	0,366	2,82
Zn	1,93	4,66
Cd	1,13	5,01
Hg	1,63	4,91
Bi	1,782	2,8
Sb	2,43	4,51

der zu fordernden Stetigkeit der Tangentialkomponenten der elektrischen und der magnetischen Feldstärke an der reflektierenden bzw. brechenden Grenzfläche zwischen zwei Medien aus. Im Ergebnis der Rechnungen für die Feldstärken der beiden linear polarisierten Komponenten gewinnt man die *Fresnelkoeffizienten*, durch die das Amplitudenverhältnis der reflektierten bzw. gebrochenen oder transmittierten Lichtanteile eindeutig festgelegt wird, wenn Einfallswinkel und optische Materialgrößen bekannt sind. Da die Intensitäten elektromagnetischer Wellen quadratisch von den Amplituden der elektrischen Feldstärke abhängen, lassen sich aus den auf Amplituden bezogenen Fresnelkoeffizienten die relativen Intensitäten unmittelbar ableiten. Man bezeichnet als *Reflektanzen* oder Reflexionsvermögen R_s und R_p die auf die einfallende Lichtintensität bezogene Intensität der reflektierten Teilwellen, als *Transmittanzen* oder Durchlässigkeiten die entsprechenden relativen Intensitäten der gebrochenen Anteile. Der Verlauf der reflektierten Anteile ist für die beiden Fälle $n_1 > n_2$ bzw. $n_1 < n_2$ in Abb. 8.8.1-4 aufgetragen. Man beachte den unterschiedlichen Gang für die senkrecht zueinander polarisierten Komponenten sowie das Auftreten von *Totalreflexion* beim Übergang aus dem optisch dichteren in das dünnere Medium.

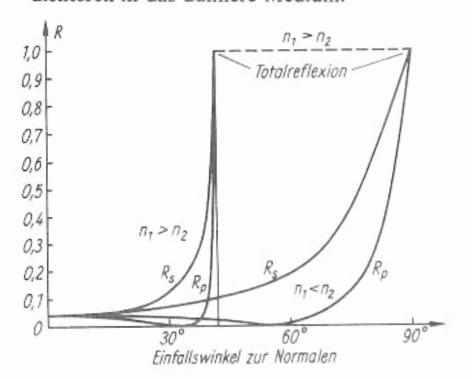


Abb. 8.8.1-4 Reflexionsvermögen R_s, R_p an der Grenzfläche von zwei Stoffen mit den Brechzahlen n_1 und n_2 für senkrecht (Index s) bzw. parallel (Index p) zur Einfallsebene polarisiertes Licht als Funktion des Einfallswinkels

Die vorstehenden Ausführungen gelten nicht nur für Festkörper, sondern allgemein für optische Medien. In Verbindung mit festen Volumen- und Schichtwerkstoffen sind die Ergebnisse jedoch von besonderer Bedeutung zur Berechnung des optischen Verhaltens beschichteter Bauteile (s. a. 8.8.1.5.).

8.8.1.4. Frequenzabhängigkeit der Materialgrößen – Dispersion. Die optischen Materialgrößen, die Brechzahl n und der Absorptionsindex a, weisen als Funktion der Frequenz bzw. der Wellenlänge des verwendeten Lichts im allgemeinen einen komplizierten Gang auf, der als *Dispersion* bezeichnet wird. Man unterscheidet dabei Bereiche normaler Dispersion, in denen n mit wachsender Frequenz bzw. abnehmender Wellenlänge steigt, von solchen anormaler Dispersion mit entgegengesetztem Frequenzgang. Die atomistische Dispersionstheorie geht davon aus, daß eine elektromagnetische Welle im Festkörper durch den Einfluß der periodisch veränderlichen elektrischen Feldstärke erzwungene Schwingungen der polarisierbaren atomaren Bausteine hervorruft. Durch die Gleichungen 8.8.(2) und 8.8.(3) werden die optischen Materialgrößen mit den elektrischen Größen Dielektrizitätskonstante und Leitfähigkeit verknüpft. Damit wird die Interpretation der optischen Dispersion auf die theoretische Beschreibung des Frequenzgangs der *elektrischen Polarisation* (↑ 8.6.2.) und der *elektrischen Leitfähigkeit* (↑ 8.6.3. und 8.3.2.) zurückgeführt. Nachfolgend sollen einige Aspekte des optischen Verhaltens für die wichtigsten Festkörperklassen qualitativ betrachtet werden.

8.8.1.4.1. *Ionenkristalle.* Typische Ionenkristalle wie NaCl sind im sichtbaren Spektralgebiet durchlässig, jedoch werden häufig, beispielsweise bei Kupfer- oder Chromiumsalzen, charakteristische Färbungen beobachtet. Die Wellenlängenabhängigkeit von n ist für einen farblosen Ionenkristall in Abb. 8.8.1-5 dargestellt. Im Infraroten steigt n mit abnehmender Wellenlänge bzw. zunehmender Frequenz zunächst flach, danach zunehmend steiler an. Dies ist auf das Anfachen transversaler optischer Phononen durch die periodisch wirkende elektrische Feldstärke der Lichtwelle zurückzuführen. Da Lichtquanten relativ nur einen sehr kleinen Impuls mitführen, werden hauptsächlich Phononen mit kleinen Wellenzahlen $q \approx 0$ angefacht, die nahezu einheitliche Frequenz besitzen (↑ Abb. 8.1.2-4). Wenn sich die Frequenz der Lichtquanten ω an diese Phononenfrequenz Ω_r annähert, so tritt Resonanz ein. In der Nachbarschaft der Resonanzfrequenz kommt es zu einer intensiven Wechselwirkung zwischen den Lichtwellen und den Gitterschwingungen. Dabei können sich gemischte Quantenzustände zwischen Photonen und Phononen ausbilden, die den Charakter von Quasiteilchen aufweisen (↑ 8.2.4.2.). Analog zu den Polaronen, an denen Elektronen und Phononen beteiligt sind, nennt man sie *Polaritonen*

Infolge anharmonischer Effekte (↑ 8.1.4.1.) wird ein Teil der Anregungsenergie von den transversalen optischen Phononen auf die akustischen und optischen Gitterschwingungen anderer Zweige umverteilt. Dies führt zu einer Dissipation von Lichtenergie in Form von Wärme.

Demnach findet in der Umgebung der Resonanzfrequenz eine starke Lichtabsorption statt. Gleichzeitig sinkt der Brechungsindex steil ab, und für $\omega > \Omega_r$ kann die Dielektrizitätskonstante formell sogar negativ werden. Physikalisch findet dann eine Totalreflexion statt, die keinen unmittelbaren Zusammenhang mit der

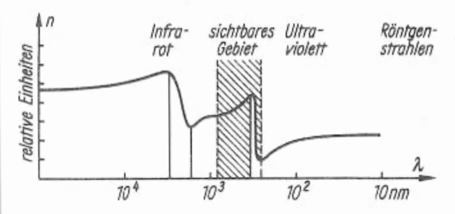


Abb. 8.8.1-5 Brechzahl n als Funktion der Lichtwellenlänge λ (schematisch); *gerastert* Bereiche anormaler Dispersion

Totalreflexion beim Übergang von optisch dichteren in dünnere Medien besitzt, sondern für den Festkörper ein „verbotenes Frequenzgebiet" verkörpert. Im Anschluß an breitere Bereiche der Absorption folgen daher bei Ionenkristallen im Infraroten schmale Frequenzgebiete einer selektiven Reflexion. Dieses Phänomen, das als Reststrahleffekt bekannt ist, wurde schon im vorigen Jahrhundert von RUBENS entdeckt, und es findet seither zum Aussondern von schmalbandigen Infrarotstrahlen aus einer kontinuierlichen Spektralverteilung Anwendung.

Mit weiter zunehmender Frequenz, d. h. im nahen Infrarotbereich, werden Ionenkristalle wieder transparent, da die Dielektrizitätskonstante ansteigt und positive Werte > 1 annimmt. Die Frequenzen optischer Phononen liegen um so höher, je kleiner die Summe der Massen der beteiligten Gitteratome sind [↑ 8.1.(13)]. Übereinstimmend damit wird beobachtet, daß Salze schwerer Ionen, beispielsweise TlBr, schon bei Wellenlängen unterhalb 50 µm durchlässig werden, während NaCl erst unterhalb 20 µm und LiF sogar erst unterhalb 8 µm Transparenz erlangen. Aus Einkristallen dieser und anderer Salze werden Fenster, Linsen, Prismen und andere Bauteile für die Infrarotoptik gefertigt.

Im sichtbaren Gebiet können die Ionen den Änderungen der elektrischen Feldstärke nicht mehr folgen. Daß der Brechungsindex dann trotzdem größer als Eins ist und überdies meist in Form der normalen Dispersion mit der Frequenz ansteigt, liegt in der Elektronenpolarisation begründet (↑ 8.6.2.). Dabei werden die äußeren Hüllenelektronen zu erzwungenen Schwingungen relativ zu den Atomrümpfen angeregt. Erneut tritt Resonanz ein, wenn die Energie der Lichtquanten $\hbar\omega$ mit dem Elektronenübergang zu einem angeregten Niveau übereinstimmt. Dies ist häufig erst im Ultravioletten der Fall, so daß viele Kristalle im gesamten sichtbaren Gebiet durchlässig und farblos sind. Fällt jedoch die Anregungsenergie für ein normalerweise unbesetztes Elektronenniveau in das sichtbare Gebiet, was beispielsweise bei den Ionen Cu^{2+} oder Cr^{3+} der Fall ist, so erfolgt in einem schmalen Frequenzintervall eine selektive Lichtabsorption, die als Färbung in Erscheinung tritt. Im Prinzip müßte danach bei höheren Frequenzen ein Gebiet mit Totalreflexion folgen, falls der Brechungsindex negative Werte annimmt. Meist tritt das nicht ein, weil der Beitrag weiterer Elektronenterme zur Polarisation ein zu starkes Absinken von ε bzw. n verhindert. Erst weit im Ultravioletten vermögen auch die Hüllenelektronen der periodisch veränderlichen Feldstärke nicht mehr zu folgen. Der Brechungsindex kann dann nur noch wenig von Eins verschieden sein, und eine effektive Lichtbrechung ist nicht mehr möglich. Dieses Verhalten, das bei allen Arten von Festkörpern beobachtet wird, bildet die Ursache dafür, daß optische Abbildungen auf der Grundlage der Lichtbrechung im entfernten Ultravioletten und erst recht im Röntgengebiet nicht ohne weiteres realisiert werden können.

8.8.1.4.2. *Molekülkristalle.* Bei Molekülkristallen können keine optischen Phononen angefacht werden, so daß nur die zuletzt besprochene Elektronenpolarisation für das optische Verhalten verantwortlich ist. Ob der Kristall im sichtbaren Gebiet transparent und farblos ist, hängt von der energetischen Lage der untersten unbesetzten Elektronenniveaus der molekularen Bausteine ab. Bei typischen Farbstoffen liegen die Quantenenergien, die zur Anregung erforderlich sind, im Bereich des sichtbaren Lichts. Manchmal sind die Resonanzen so ausgeprägt, beispielsweise bei metallisch glänzenden Farbstoffkristallen von Eosin oder Fluorescein, daß oberhalb der Anregungsfrequenz ein Absinken der Dielektrizitätskonstante auf Werte unter Null und damit eine Zone mit Totalreflexion auftritt.

In einigen Molekülkristallen organischer Stoffe sind auch erzwungene Schwingungen von Dipolgruppen nicht zu vernachlässigen, obwohl dieser Einfluß aus sterischen Gründen geringer ist als bei Festkörpern oder Gasen. Auf jeden Fall erlischt der Dipolbeitrag schon bei relativ großen Wellenlängen, also im fernen Infrarot oder gar im Mikrowellengebiet.

8.8.1.4.3. *Metalle.* Die auffälligen optischen Eigenschaften der Metalle, d. h. nahezu Totalreflexion und sehr starke Lichtabsorption, sind auf das Verhalten der leicht verschieblichen quasifreien Elektronen in der Nähe der *Fermi-fläche* zurückzuführen (↑ 8.3.3., 8.3.4.). Diese Leitungselektronen führen unter dem Einfluß der elektrischen Feldstärke der Lichtwelle periodische Bewegungen aus, wobei sie dem Ohm-

schen Widerstand entsprechend, hauptsächlich durch Elektron-Phonon-Wechselwirkung, laufend Energie an das Gitter abgeben, was natürlich auf Kosten der Lichtenergie geschieht. Infolgedessen absorbieren Metalle sehr stark, und sie verhalten sich optisch etwa wie Molekülkristalle, bei denen für jede Frequenz die Resonanzbedingung erfüllt ist. Der Brechungsindex verbleibt bis zum ferneren Ultraviolettbereich im Negativen, woraus die starke Reflexion folgt, die schon seit ältesten Zeiten für die Herstellung von Spiegeln genutzt wird. Erst weit im Ultravioletten bei Frequenzen oberhalb der Plasmonenfrequenz (↑ 8.2.4.2.) $\omega_p^2 = (ne^2)/(\varepsilon_0 m^*)$ können die Metallelektronen dem anregenden elektrischen Feld der Lichtwellen nicht mehr folgen, und die Metalle erlangen dann zunehmende Transparenz. Im übrigen ist zur feineren Deutung der optischen Eigenschaften von Metallen die Kenntnis der Fermiflächen erforderlich.

8.8.1.4.4. *Halbleiter.* Die Valenzkristalle und speziell die Halbleiter nehmen in ihrem optischen Verhalten eine Zwischenstellung zwischen Metallen und Molekülkristallen ein. Im Infraroten sind sie transparent, was beispielsweise in IR-Mikroskopen zur Durchsichtsbeobachtung von Ge oder Si genutzt wird. Dagegen erfolgt im sichtbaren Gebiet und schon etwas früher eine starke Absorption, die mit einem schwach metallischen Glanz einhergeht. Maßgeblich für den Übergang zur Absorption ist der Bandabstand E_g (↑ 8.3.1.). Solange die Energie der Lichtquanten nicht ausreicht, um Elektronen über die Bandlücke vom Valenzband in das Leitungsband anzuheben, bleibt die Absorption gering, und der Brechungsindex wird von der Elektronenpolarisation bestimmt. Bei höheren Frequenzen werden Elektronen von Band zu Band angeregt und nachfolgende, häufig strahlungslose Rekombinationsprozesse bewirken eine Dissipation der Lichtenergie (↑ 8.3.4.3.). Im einzelnen liegen die Verhältnisse freilich komplizierter (↑ 8.8.3.), da auch die Beteiligung von Störstellenniveaus zu berücksichtigen ist. Andererseits wird vielfach schon bei Quantenenergien unterhalb des Bandabstands eine verstärkte Lichtabsorption festgestellt, weil sich bei unvollständiger Trennung von Elektronen und Löchern *Exzitonen* bilden können (↑ 8.2.4.2.2.).

8.8.1.5. Optische Schichtsysteme. Für Bauteile optischer Geräte, aber auch schon für einfache Spiegel und Fenster ist es wünschenswert oder sogar erforderlich, daß Oberflächen in einem bestimmten Spektralbereich reflektierende oder umgekehrt antiflektierende Eigenschaften aufweisen. So werden Spiegel seit Jahrhunderten mit Hilfe von Metallschichten, ursprünglich Hg, später Ag und jetzt Al, hergestellt, die man chemisch oder gegenwärtig durch Aufdampfen oder Aufstäuben im Vakuum auf Glas abscheidet. Mit den immer größer werdenden Glasfronten klimatisierter Gebäude entstand aus Gründen der Energieökonomie ein Bedarf an Bauglas, das im sichtbaren Spektralbereich gut durchlässig, im Infraroten dagegen reflektierend sein soll. Auch dies läßt sich durch Aufbringen von Schichten im Vakuum lösen. Man verwendet heute meist eine sehr dünne Metallschicht, die zwischen transparenten SiO_x- oder TiO_2-Schichten eingebettet ist. Auf die gleiche Weise lassen sich auch Glühlampen realisieren, bei denen durch Reflexion des nutzlosen Infrarotlichts in das Innere ein nahezu um das Doppelte erhöhter energetischer Wirkungsgrad möglich erscheint.

Kompliziertere Anforderungen sind bei den Vergütungsschichten optischer Präzisionsbauteile zu erfüllen. Normalerweise werden an Glasoberflächen selbst bei kleinem Einfallswinkel mehrere Prozent des auffallenden Lichts reflektiert (↑ Abb. 8.8.1-1). Bei optischen Geräten wie Kameras, Teleskopen, Mikroskopen ginge dadurch ein erheblicher Teil des eintretenden Lichts für die Abbildung verloren, oder, was noch schlimmer wäre, es würden durch nachfolgende Reflexionen „Geisterbilder" erscheinen. Noch strengere Anforderungen stellen medizinische Geräte zur Beobachtung des Körperinneren (Zystoskope) oder die Sichtgeräte von U-Booten (Periskope), die bis etwa 40 aufeinanderfolgende brechende Glasflächen enthalten können.

Die Vergütung der Glasoberflächen aller wichtigen optischen Bauteile erfolgt heute durch das Aufbringen von Schichtfolgen mit alternierender Brechzahl bei genau einzuhaltender Dicke. Bereits durch Aufdampfen einer einzigen Schicht mit einer kleineren Brechzahl als Glas läßt sich die Reflektanz bzw. das Reflexionsvermögen herabsetzen (Abb. 8.8.1-6). Der Effekt ist optimal, wenn sich die Teilstrahlen, die an der äußeren und der inneren Schichtoberfläche reflektiert werden, im Lichtweg gerade um eine halbe Wellenlänge unterscheiden, die Schichtdicke,

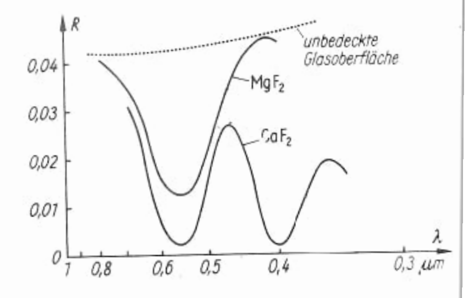


Abb. 8.8.1-6 Änderung des Reflexionsvermögens R einer Glasoberfläche durch Aufbringen dünner Schichten

also $\lambda/4$ entspricht. Durch Interferenz der reflektierten Teilwellen kann dann Auslöschung eintreten, und es ist keine Reflexion möglich. Genauere Untersuchungen zeigen indessen, daß eine vollständige Auslöschung den Einsatz von Schichtmaterialien mit sehr kleiner Brechzahl erfordern würde. Da solche Materialien nicht verfügbar sind, bringt man zuerst eine Schicht mit einer gegenüber Glas erhöhten Brechzahl und danach einen Film mit kleiner Brechzahl auf. Eine weitere Schwierigkeit besteht darin, daß die Auslöschung durch Interferenz zunächst nur für eine ganz bestimmte Wellenlänge und das unmittelbar benachbarte Gebiet eintritt. Um diesen Bereich zu verbreitern, verwendet man vielfache Stapelfolgen aus bis zu 20 Schichten mit abwechselnd hoher bzw. niedriger Brechzahl. Daß dabei eine Verbreiterung des nahezu reflexfreien Frequenzgebiets herauskommt, beruht auf den gleichen Gesetzmäßigkeiten der Wellenausbreitung wie die dem Funkamateur bekannte Wirkung von Bandpaßfiltern oder auch die vielfache Aufspaltung der Eigenwerte in gekoppelten atomaren Systemen (↑ 8.2.1.2.2.). In einfacheren Fällen kann man schon mit vier Schichten eine weitgehende Senkung der Reflexion im sichtbaren Gebiet erzielen. Bei höherem Aufwand mit zehn und mehr Schichten läßt sich die Reflektanz im gesamten sichtbaren Bereich kleiner als 0,002 halten. Als Materialien mit niedriger Brechzahl werden vorzugsweise MgF_2, SiO_2 oder TiO_2 eingesetzt, während ZnS, GeO_2 oder Si_3N_4 für die Filme höherer Brechzahl Verwendung finden.

Auf dem gleichen Prinzip und unter Benutzung der gleichen Materialien geschieht die Herstellung von nahezu idealen Spiegeln mit einer Reflektanz von 0,998, wie sie vor allem für die Laseroptik benötigt werden. Jede der etwa acht bis zwölf Schichten besteht aus durchsichtigem Material, jedoch wird durch Einhaltung der optischen Schichtdicke $\lambda/2$ erreicht, daß sich die transmittierten Strahlen auslöschen und das gesamte Licht in einem definierten Spektralgebiet reflektiert werden muß. Unvermeidliche Verluste, die durch ständige Verbesserung der Beschichtungsverfahren weiter gesenkt werden, treten durch ungenügend beherrschte Toleranzen und vor allem durch Lichtstreuung bzw. -absorption an Schichtbaufehlern auf.

Eine wichtige neuere Anwendung reflektierender Schichtfolgen bilden die hohlen Glasfasern für die Lichtleiternachrichtenübertragung. Man beschichtet Glasrohre auf der Innenseite mit einer Vielfachschicht und zieht die Rohre danach gleichmäßig zu feinen Kapillaren aus. Die Übertragungsgüte der daraus gefertigten *Lichtleiterkabel* wird maßgeblich von der Qualität des reflektierenden Schichtsystems bestimmt.

Die zunehmende Beherrschung von Mikrodimensionen bei optischen Bauteilen hat den Anstoß gegeben, analog zur Mikroelektronik eine Art integrierter Optik zu entwickeln. Obwohl der Stand vorläufig nicht mit dem der Mikroelektronik (↑ 8.3.7.) verglichen werden kann, wurden bei Systemen passiver optischer Bauelemente schon beachtliche Erfolge erzielt. Im Verlauf der künftigen Entwicklung wird man versuchen, unter Einbeziehung geeigneter aktiver Elemente, besonders auf der Basis von Lasern, ein System optischer Informationsverarbeitung zu schaffen, für das der Name *Optronik* geprägt worden ist.

8.8.2. Anisotropie und nichtlineare Optik

Kristalle, die nicht dem kubischen System angehören, verhalten sich optisch anisotrop, d. h., die optischen Eigenschaften sind richtungsabhängig. Aber auch isotrope Stoffe, darunter Gläser und Plaste, können optische Anisotropie erlangen, wenn man durch starke elektrische oder magnetische Felder eine Vorzugsrichtung schafft. Das optisch anisotrope Verhalten tritt besonders auffällig als Aufspaltung von Licht in zwei Teilstrahlen in Erscheinung. Man beobachtet dies beispielsweise an Kalkspatplatten als Doppelbrechung, bei der ein in Durchsicht betrachtetes Bild in zwei zueinander verschobene Teilbilder aufspaltet. Die Erforschung der optischen Anisotropie hat in der Geschichte der Kristallographie eine wichtige Rolle gespielt. Praktische Anwendungen findet die Anisotropie bei der Erzeugung und Analyse von polarisiertem Licht. Neuerdings haben optisch anisotrope Kristalle darüber hinaus in der nichtlinearen Optik für die Erzeugung von Oberwellen besonderes Interesse erlangt.

8.8.2.1. Besonderheiten optisch anisotroper Kristalle. Im Falle optischer Anisotropie hängt die Phasengeschwindigkeit des Lichts im Festkörper von der Ausbreitungsrichtung ab, und zu jeder Richtung existieren im allgemeinen zwei zulässige Werte. Als Folge davon spalten Lichtwellen beim Eintritt in den Kristall in zwei Teilstrahlen auf, die senkrecht zueinander polarisiert sind, bei denen also die elektrische Feldstärke in zwei zueinander senkrechten Ebenen schwingt (↑ 2.3.1.). Da die Brechzahl gleich dem Verhältnis zwischen Vakuumlichtgeschwindigkeit und Phasengeschwindigkeit des Lichts im Medium ist, werden die beiden Teilstrahlen unterschiedlich gebrochen, so daß sie sich normalerweise in verschiedenen Richtungen ausbreiten. Längs bestimmter Vorzugsrichtungen, die man als *optische Achsen* bezeichnet, breiten sich die Teilstrahlen jedoch einheitlich aus, und es kommt nicht zur Aufspaltung des Lichts. Bei den Kristallen des rhomboedrischen, tetragonalen oder hexagonalen Systems gibt es *eine* solche Vorzugsrichtung, die mit einer kristallographi-

schen Achse zusammenfällt; diese Kristalle sind optisch einachsig. Die weniger symmetrischen Kristalle des rhombischen, monoklinen und triklinen Systems besitzen dagegen zwei optische Achsen, die im Falle monokliner und trikliner Kristalle auch nicht senkrecht aufeinander stehen. Außerdem können die Achsenrichtungen bei Kristallen geringster Symmetrie selbst von der Frequenz des Lichts abhängen. Man spricht dann von *Achsendispersion.*

Der tiefere Grund für die optische Anisotropie ist darin zu suchen, daß die vom elektrischen Feld der Lichtwelle erzeugte Polarisation in den verschiedenen Kristallrichtungen unterschiedlich stark erfolgt. Bei nicht zu hohen Feldstärken hängen die Vektoren der Polarisation und der Verschiebung gemäß $\boldsymbol{P} = \varepsilon_0 \varkappa_{el} \boldsymbol{E}$ bzw. $\boldsymbol{D} = \varepsilon_0 \varepsilon \boldsymbol{E}$ auch im Falle der Anisotropie linear von der elektrischen Feldstärke E ab. Allerdings sind die elektrische Suszeptibilität $\varkappa_{el}$ und die Dielektrizitätskonstante ε dann keine skalaren Größen, sondern symmetrische Tensoren mit maximal sechs unabhängigen Komponenten (↑ 8.6.2.). Die obigen Linearitätsbeziehungen werden somit lineare Vektorfunktionen, und $\boldsymbol{P}$ bzw. $\boldsymbol{D}$ sind nicht länger mit $\boldsymbol{E}$ gleichgerichtet, außer in Vorzugsrichtungen.

Als symmetrischer Tensor läßt sich ε – und analog auch $\varkappa_{el}$ – auf Hauptachsen transformieren. Auf diese Weise erhält man die drei Hauptdielektrizitätskonstanten, die sich in der geometrischen Darstellung des Indexellipsoids als Halbachsen veranschaulichen lassen (Abb. 8.8.2-1). Man betrachtet nur die Ausbreitung von Wellen, die sich vom Ursprung des Hauptachsensystems ξ, η, ζ aus fortpflanzen, wobei der Einheitsvektor s_n^0 die Richtung senkrecht zu den Wellenfronten, die sogenannte Normalenrichtung angeben soll. Ebenen, die senkrecht zu s_n^0 durch den Koordinatenursprung errichtet werden, schneiden das Ellipsoid normalerweise in elliptischen Schnittflächen. Die beiden Hauptachsen dieser Schnittellipsen geben die Richtungen $\boldsymbol{D}_1$ und $\boldsymbol{D}_2$ an, in denen die elektrische Verschiebung der beiden senkrecht zueinander polarisierten Teilwellen schwingt. Die Quadrate der zugehörigen Brechungsindizes sind wegen des allgemeinen Zusammenhangs $n^2 = (c/v)^2 = \varepsilon$ gleich halben Achsen bzw. Durchmessern dieser Schnittellipsen.

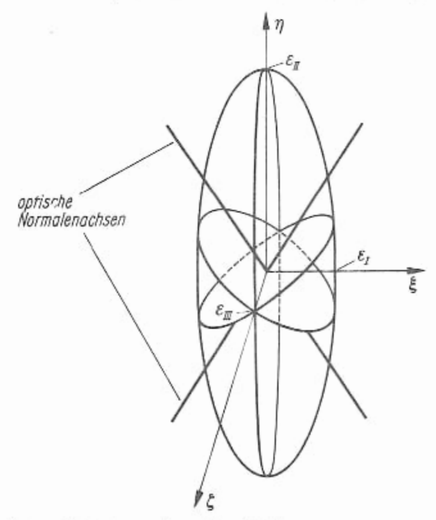


Abb. 8.8.2-1 Indexellipsoid im Hauptachsensystem ξ, η, ζ, sowie optische Achsen eines zweiachsigen Kristalls

Bei zweiachsigen Kristallen gibt es nun genau zwei Richtungen, deren Normalebenen das Indexellipsoid in Form von Kreisen schneiden. Die Brechungsindizes oder Brechzahlen der Teilwellen sind für diese Richtungen gleich, und es findet keine Aufspaltung statt. Diese Richtungen heißen *optische Normalenachsen.* Sie spannen die Ebene der optischen Achsen auf, und die Winkelhalbierende zwischen ihnen wird *optische Mittellinie* genannt. Eine weitere Komplikation tritt auf, weil die mit s_n^0 eingeführte Richtung der Normalen auf den Wellenfronten konstanter Fläche nicht mit der Strahlrichtung s_s^0 übereinstimmt, in der sich der Energiestrom ausbreitet. Zwischen Normalen- und Strahlrichtung gilt die Beziehung $s_s^0 \cdot s_n^0 = \cos \alpha$, d. h., die Wellenfronten konstanter Phase sind unter dem Winkel α zur Strahlrichtung geneigt. Dabei schwingt die elektrische Verschiebung $\boldsymbol{D}$ senkrecht zur Normalenrichtung, die Feldstärke $\boldsymbol{E}$ dagegen transversal zur Strahlrichtung. Der richtungsabhängige Brechungsindex $n = c/v_n$, der für jede Richtung im allgemeinen zwei Werte annehmen kann, stellt das Verhältnis zwischen Vakuumlichtgeschwindigkeit c und Phasengeschwindigkeit v_n in Normalenrichtung dar, während die Strahlgeschwindigkeit, mit der sich die Energie in Richtung s_s^0 fortpflanzt, $v_s = v_n/\cos \alpha$ beträgt.

Anschauliche Informationen über die Wellenausbreitung in anisotropen Festkörpern werden durch die Konstruktion der Strahlen- und Normalenflächen erhalten, indem vom Koordinatennullpunkt aus die Strahlgeschwindigkeit v_s bzw. die Normalengeschwindigkeiten v_n aufgetragen werden, die den Richtungen s_s^0 bzw. s_n^0 zugeordnet sind. Die Flächen stellen somit den geometrischen Ort der energetischen Lichterregung bzw. der Phase aller Wellen dar, die eine Zeiteinheit zuvor im Nullpunkt des Hauptachsensystems gestartet sind. Bei zweiachsigen Kristallen bestehen die Strahlen- bzw. Normalenflächen aus je zwei geschlossenen Schalen, die an den vier Durchstoßpunkten der optischen Achsen zusammenhängen.

8.8.2.2. Eigenschaften optisch einachsiger Kristalle. Erhebliche Vereinfachungen ergeben sich bei den optisch einachsigen Kristallen, zu denen Kalkspat $CaCO_3$ und (α)-Quarz SiO_2 gehören. Das Indexellipsoid geht in ein Rotationsellipsoid über, da zwei der drei Hauptdielektrizitätskonstanten einander gleich sind. Mit $n_{or}^2 = \varepsilon_{or} = \varepsilon_I = \varepsilon_{II}$ und $n_{ao}^2 = \varepsilon_{ao} = \varepsilon_{III}$ führt man die beiden Brechzahlen des ordentlichen und des außerordentlichen Strahls ein. Diese historisch bedingten Bezeichnungen rühren daher, daß der *ordentliche Strahl*, der senkrecht zu dem aus Einfallsrichtung und optischer Achse aufgespannten Hauptschnitt schwingt, unabhängig von der Kristallorientierung gebrochen wird. Der *außerordentliche Strahl*, bei dem die elektrische Feldstärke parallel zum Hauptschnitt polarisiert ist, erfährt dagegen eine von der Kristallorientierung abhängige Brechung. (Bei zweiachsigen Kristallen verhalten sich beide Teilstrahlen außerordentlich.) Die Strahlenflächen optisch einachsiger Kristalle bestehen aus einem Rotationsellipsoid und einer Kugelfläche, die sich an den beiden Durchstoßpunkten der optischen Achse berühren (Abb. 8.8.2-2). Man unterscheidet dabei zwischen einachsig negativen oder positiven Kristallen, je nachdem n_{ao} kleiner oder größer als n_{or} ist. Für Anwendungen eignet sich besonders der stark doppeltbrechende Kalkspat mit einem Verhältnis $n_{or}/n_{ao} = 1{,}1157$, von dem große Einkristalle guter optischer Qualität in der Natur gefunden werden. Zur Erzeugung und Analyse von linear polarisiertem Licht benutzt man vielfach die *Nicolschen Prismen* (Abb. 8.8.2-3). Zu ihrer Herstellung wird ein passend geschliffenes Kristallstück diagonal zersägt, und die Teilstücke fügt man mit Klebstoff wieder zusammen. Bei richtiger Geometrie wird der ordentliche Strahl durch Totalreflexion an der Grenze zwischen dem optisch dichteren Kristall und dem Kleber abgelenkt, während der außerordentliche Strahl durch die Grenzflächen hindurchtritt. Mit einem zweiten Prisma dieser Art läßt sich die Polarisationsrichtung nachweisen, indem man die Prismen relativ zueinander um die Strahlrichtung dreht. So erscheint eine Lichtquelle dunkel, wenn man sie durch „gekreuzte", d. h. um 90° gegeneinander gedrehte Nicols beobachtet. Das erste Prisma sondert dabei aus den regellos orientierten Wellenzügen diejenigen Komponenten aus, die gemäß der Durchlaßrichtung polarisiert sind. Im zweiten Prisma entspricht diese Polarisation jedoch gerade dem totalreflektierten Teilstrahl, so daß längs der Achse kein Licht austritt.

8.8.2.3. Optische Aktivität. Bringt man Kristalle, die kein Inversionszentrum und keine Spiegelebenen (↑ 6.4.1.) besitzen, zwischen gekreuzte Nicolprismen, so wird eine Aufhellung festgestellt. Die gleiche Eigenschaft zeigen Lösungen organischer Stoffe, z. B. Traubenzucker, deren Moleküle ein asymmetrisch gebundenes Kohlenstoffatom besitzen (↑ 5.2.1.). Dieses charakteristische Verhalten, das als optische Aktivität bezeichnet wird, beruht auf einer Drehung der Polarisationsebene von linear polarisiertem Licht. Der Drehwinkel ergibt sich aus dem Pro-

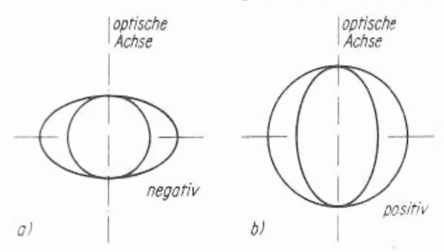


Abb. 8.8.2-2 Schnitt durch die Strahlenflächen a) eines einachsig negativen, b) eines einachsig positiven Kristalls; die Strahlenflächen einachsiger Kristalle sind Rotationsellipsoide bzw. Kugelflächen

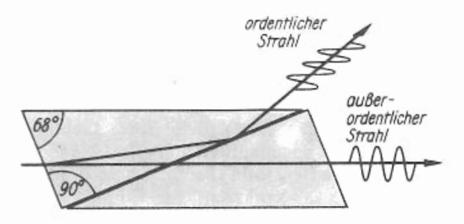


Abb. 8.8.2-3 Erzeugung von linear polarisiertem Licht durch Trennung des ordentlichen Strahls vom außerordentlichen Strahl im Nicolschen Prisma

dukt des Lichtwegs im Medium und der spezifischen Drehung α_D, die eine charakteristische, allerdings frequenzabhängige Materialgröße bildet. Beispielsweise drehen Quarzscheiben, die senkrecht zur optischen Achse geschnitten sind, die Polarisationsrichtung um 29,72° pro mm Dicke bei einer Lichtwellenlänge von 508,6 nm. Die optische Aktivität überlagert sich bei Kristallen häufig der Anisotropie. Die Strahlenflächen optisch aktiver Stoffe berühren sich nicht mehr, sondern sind voneinander getrennt. Sogar einige kubische Kristalle wie $NaClO_3$ zeigen optische Aktivität. Ihre Strahlenflächen bestehen aus zwei konzentrisch angeordneten Kugelschalen. In der chemischen Analytik wird die Messung der spezifischen Drehung mit Hilfe von Polarimetern häufig zur Konzentrationsbestimmung herangezogen. Auch in der Mikroskopie werden Polarisatoren häufig eingesetzt, um den Kontrast durch vielfach farbenprächtige Interferenzen in polarisiertem Licht zu verbessern.

8.8.2.4. Dichroismus, Polarisationsfilter. Eine

andere Möglichkeit zur Erzeugung von polarisiertem Licht bildet die bei einigen Stoffen beobachtete unterschiedliche Absorption der senkrecht zueinander polarisierten Teilstrahlen. Man bezeichnet diesen Effekt als Dichroismus. Besonders ausgeprägt dichroitisch verhalten sich die Turmaline, das sind Aluminiumborosilikate komplizierter Zusammensetzung, die als Mineralien in der Natur vorkommen und auch als Halbedelsteine bekannt sind. Eine Turmalinscheibe von 1 mm Dicke, die parallel zur optischen Achse aus einem Kristall herausgesägt ist, absorbiert den ordentlichen Teilstrahl fast völlig, während der außerordentliche Strahl noch in brauchbarer Intensität durchgelassen wird. Eine moderne Entwicklung bilden stark dichroitische Polymerfolien, die man durch Verstrecken erhält, wobei sich die Kettenmoleküle ausrichten. Diese nahezu farblosen *Polaroidfolien* lassen sich auch in großen Abmessungen herstellen und gewinnen als Polarisatoren wachsende Bedeutung.

Der Vollständigkeit halber sei darauf hingewiesen, daß auch die *Reflexion* an einer Grenzfläche zur Erzeugung von polarisiertem Licht benutzt werden kann. Dies beruht auf der unterschiedlichen Reflexion der senkrecht bzw. parallel zur Einfallsebene polarisierten Teilwellen (↑ 8.8.1.3., Abb. 8.8.1-4). Erfolgt der Lichteinfall unter dem sogenannten *Brewsterwinkel*, so ist der reflektierte Strahl theoretisch vollständig linear, und zwar senkrecht zur Einfallsebene, polarisiert. Praktisch ist das wegen Inhomogenitäten der Oberfläche nicht der Fall, doch wird die Polarisation durch Reflexion, z. B. an Selen, im Infraroten genutzt.

Die leistungsfähigsten Quellen für polarisiertes Licht, das außerdem kohärent ist, bilden gegenwärtig die *Laser* (↑ 4.1.4.). Um den Aufbau des linear polarisierten Wellenfelds in Lasern zu begünstigen, werden die teildurchlässigen Endspiegel vielfach unter dem Brewsterwinkel angesetzt.

8.8.2.5. Nichtlineare Optik und Phasenanpassung. Die Entwicklung leistungsstarker Laser eröffnete den experimentellen Zugang zu Erscheinungen der Wechselwirkung von Licht mit Festkörpern, bei denen nichtlineare Prozesse eine wesentliche Rolle spielen. Aufbauend auf theoretische Vorarbeiten von WAWILOW wurde die nichtlineare Optik vor allem durch ACHMANOW und CHOCHLOW in der UdSSR und durch FRANKEN, BLOEMBERGEN und DUOUING in den USA um die Mitte der 60er Jahre begründet.

Bei genügend hoher elektrischer Feldstärke der Lichtwelle besteht zwischen der elektrischen Polarisation P und der Feldstärke E kein linearer Zusammenhang mehr, sondern die elektrische Suszeptibilität $\varkappa_{\text{el}}$ wird selbst eine Funktion der Feldstärke. In einer Dimension läßt sich die Beziehung zwischen P und E durch Entwicklung in eine Potenzreihe näherungsweise mit

$$(4) \quad P = \varkappa_{\text{el}}^{(1)}E + \varkappa_{\text{el}}^{(2)}E^2 + \varkappa_{\text{el}}^{(3)}E^3$$

wiedergeben. Bei einachsigen Kristallen ohne Symmetriezentrum, z. B. Quarz, überwiegt das quadratische Glied gegenüber dem kubischen, während bei Vorhandensein eines Symmetriezentrums $\varkappa_{\text{el}}^{(2)}$ verschwindet und nur das kubische Glied einen Beitrag liefert. Ähnlich wie bei den früher betrachteten *anharmonischen Effekten* der Gitterschwingungen (↑ 8.1.4.1.) bewirkt die Nichtlinearität, daß die Wellen, hier also die Photonen, nicht mehr voneinander unabhängig sind, sondern daß eine *Photon-Photon-Wechselwirkung* erfolgt. Sind zunächst nur Photonen einheitlicher Frequenz vorhanden, so kommt es zur Frequenzverdopplung bzw. -verdreifachung oder auch umgekehrt zur Frequenzteilung. Bei Photonenbündeln unterschiedlicher Frequenz können sich ferner Wellen der Summen- oder Differenzfrequenzen ausbilden. Experimentell beobachtete FRANKEN 1961 erstmalig eine Frequenzverdopplung, als er Quarzscheiben der Strahlung eines Rubinlasers aussetzte.

Allerdings ist die Effektivität bei der Bildung von Oberwellen normalerweise gering. Zwar nimmt die Intensität der Wellen doppelter Frequenz zunächst längs des Lichtwegs im Kristall zu, jedoch erfolgt danach ein Abflachen des Wachstums und schließlich sinkt die Intensität der Oberwelle wieder auf Null ab. Danach erfolgt ein neues Wachstum usw. Der Abstand zwischen benachbarten Nullstellen ist durch die *Kohärenzlänge* $k \to l_k = \lambda/4(n_2 - n_1)$ gegeben, wenn λ die Wellenlänge der Grundwelle und n_2 bzw. n_1 die Brechzahlen der Wellen mit der doppelten Frequenz 2ω bzw. der Grundfrequenz ω bezeichnen. Eine optimale Intensität der Oberwelle ließe sich erzielen, wenn die Bedingung $n_2 = n_1$ erfüllt werden könnte. Grund- und Oberwelle breiten sich dann mit gleicher Ge-

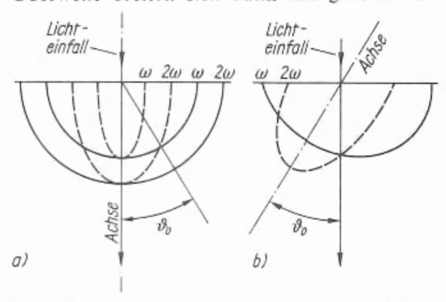


Abb. 8.8.2-4 Phasenanpassung von Oberwellen; *gestrichelte Linien* Strahlenflächen der außerordentlichen, *durchgezogene Linien* der ordentlichen Lichtwellen mit den Frequenzen ω bzw. 2ω; a) Ausbreitung parallel zur optischen Achse, b) in Anpassungsrichtung

schwindigkeit aus, und die Kohärenzlänge wird – wenigstens theoretisch – unendlich. Praktisch ist dies bei isotropen Kristallen wegen der Dispersion nicht zu erfüllen. Einen Ausweg eröffnet der Einsatz von anisotropen Kristallen, da man sich dann Richtungen aussuchen kann, für welche der Brechungsindex des außerordentlichen Strahls der Frequenz 2ω mit dem des ordentlichen Strahls der Frequenz ω übereinstimmt. Abb. 8.8.2-4 veranschaulicht das

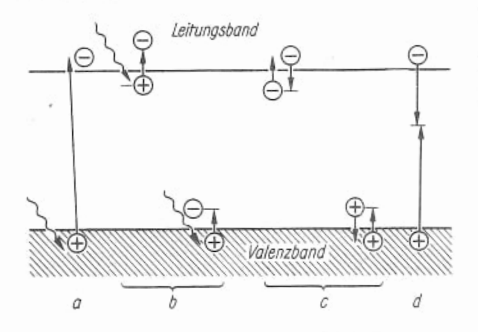


Abb. 8.8.3-1 Elementarprozesse bei der Fotoleitung in Kristallen; *a* Lichtabsorption im Grundgitter, *b* Befreiung von Elektronen bzw. Defektelektronen aus nichtionisierten Störstellen, *c* Einfang und Freisetzung von Ladungsträgern an Haftstellen, *d* Rekombination nach SHOCKLEY und READ

darauf basierende Prinzip der *Phasenanpassung*. Die Verhältnisse entsprechen dem einachsig negativen Kaliumhydrogensulfat KH_2PO_4, meist kurz als KDP bezeichnet, das besonders häufig für die Frequenzverdopplung von Laserlicht eingesetzt wird.
Im Teilbild a sind die Strahlenflächen für die beiden Wellen im Schnitt dargestellt. Man erkennt, daß die Phasengeschwindigkeiten und damit auch die Brechungsindizes auf einem Kegel mit dem Öffnungswinkel $2\vartheta_0$ gleich sind. Um eine effektive Frequenzverdopplung zu erzielen, schneidet man Platten so aus dem Kristall heraus, daß ein Strahl dieses Kegels senkrecht zur Oberfläche steht. Die optische Achse schließt dann mit der Oberflächennormalen den Winkel ϑ_0 ein. Läßt man nun Laserlicht mit einer Polarisation entsprechend dem ordentlichen Strahl senkrecht in die Platte einfallen, so kommt es bei gleicher Geschwindigkeit der Teilwellen im Kristall zu einer intensiven und anhaltenden Frequenzverdopplung. Praktisch werden heute Ausbeuten von mehr als 20% an Oberwellen bezogen auf die einfallende Intensität erreicht. Analog läßt sich bei Verwendung von Kristallen mit Symmetriezentrum, z. B. Kalkspat, eine Frequenzverdreifachung bei allerdings geringerer Ausbeute durchführen. Zur Phasenanpassung muß man dann Kristallplatten in der Weise schneiden, daß die Brechzahlen bzw. Phasengeschwindigkeiten der Wellen mit den Frequenzen ω und 3ω gleich werden.
Die Frequenzvervielfachung sowie die analoge Frequenzteilung werden heute vielfältig genutzt, um den Frequenzbereich von Laserstrahlen zu erweitern. Darüber hinaus zielen weitere Anwendungen der nichtlinearen Optik auf die Entwicklung von parametrischen Lichtgeneratoren und Lichtverstärkern ab. In Verbindung mit integrierten optischen Bauelementen kann sich daraus künftig eine Informationsverarbeitung auf der Basis optischer Systeme in Gestalt einer Optronik ergeben (8.8.1.5.).

8.8.3. Fotoeffekte – Optoelektronik

Als Fotoeffekte im weiteren Sinn bezeichnet man die Erscheinungen in Festkörpern, die durch Wechselwirkung zwischen Elektronen und Photonen, häufig unter Beteiligung von Phononen und anderen Quasiteilchen, hervorgerufen werden. Besonders bedeutungsvoll sind die durch Lichteinstrahlung bewirkte Fotoleitung sowie umgekehrt die zur Lichtemission führenden elektrischen Leitungsprozesse, ferner auch die Erscheinungen der Lumineszenz und Phosphoreszenz. Die Erforschung und Nutzung der Fotoeffekte hat in den letzten Jahren einen bedeutenden Aufschwung genommen. Als Triebkraft wirkt vor allem die Ergänzung der Mikroelektronik durch eine Optoelektronik, die auf einer kombinierten Anwendung elektrischer und optischer Vorgänge in Festkörpern beruht.
8.8.3.1. Fotoleitung. In Kristallen von Halbleitern und Isolatoren werden bei der Absorption genügend energiereicher Lichtquanten verschiedenartige Elementarprozesse ausgelöst (Abb. 8.8.3-1). Wenn die Energie der Lichtquanten $\hbar\omega$ größer ist als der Bandabstand (a), so können Elektronen aus dem Valenzband in das Leitungsband angeregt werden. Dabei entstehen paarweise Ladungsträger in Form von Elektronen und Löchern, die bei Anlegen einer Spannung zu einer erhöhten Leitfähigkeit Anlaß geben (↑ 8.3.4.1.). Falls die Quantenausbeute, d. h. die Anzahl der angeregten Elektronen, gleich Eins ist, so werden durch z Lichtquanten $2z$ zusätzliche Ladungsträger gebildet. Weiterhin können Ladungsträger auch durch optische Befreiung von Ladungsträgern aus Störstellen gebildet werden (b). Die hierfür erforderliche Quantenenergie ist naturgemäß kleiner als bei der Bildung von Trägerpaaren durch Interbandübergänge. Als Haftstellen (c) werden lokalisierte Elektronenzustände im Festkörper, beispielsweise an Störstellen oder Oberflächenzuständen, bezeichnet,

in denen Elektronen oder Löcher eingefangen und vorübergehend festgehalten werden. Die eingefangenen Ladungsträger nehmen zunächst an der Fotoleitung nicht mehr teil; sie werden jedoch meist nach kurzer Zeit durch thermische oder optische Anregung wieder befreit. Der zusätzlichen Bildung von Ladungsträgern bei Belichtung wirkt die *Rekombination* (d) entgegen (↑ 8.3.4.2.). Wird die Belichtung des Kristalls abgebrochen, so klingt die zusätzliche Ladungsträgerkonzentration innerhalb von 10^{-6} bis 10^{-1} s wieder ab. Mit Hilfe von Lichtblitzmethoden lassen sich die Trägerlebensdauern zusätzlicher Elektronen und Löcher in Halbleitern bestimmen.

Das hier entworfene Bild der Generation von Ladungsträgern durch Licht ist in vieler Hinsich noch sehr grob. Eine genauere Betrachtung anhand des $E(k)$-Diagramms der Energiebänder (↑ 8.2.1.) lehrt, daß insbesondere bei indirekten Halbleitern häufig noch die Mitwirkung von Phononen erforderlich ist, um bei den Interbandübergängen gleichzeitig den Energie- und den Impulserhaltungssatz zu erfüllen. Die Phononen liefern dabei den Impulsanteil, der zum Übergang der Elektronen in Zustände unter wesentlicher Änderung des Wellenzahlvektors $\boldsymbol{k}$ benötigt wird (Abb. 8.8.3-2). Andererseits wird eine verstärkte Lichtabsorption in Halbleitern schon beobachtet, wenn die Quantenenergie etwas kleiner ist als der Bandabstand. Ursache hierfür ist die Bildung von Exzitonen, d. h. aneinander gebundener Elektron-Lochpaare (↑ 8.2.4.2.2.). Diese Exzitonen tragen zwar nicht zum Stromfluß bei, da sie als Ganzes gesehen elektrisch neutral sind, sie können jedoch durch das Gitter diffundieren und an nichtbelichteten Stellen des Kristalls ihre Anregungsenergie abgeben, indem sie rekombinieren. Die Bildung von Exzitonen macht verschiedenartige Effekte verständlich, und sie ist eine Ursache, warum die Absorption von Lichtquanten nicht immer eine Fotoleitung zur Folge hat.

Der Stromfluß in Fotoleitern hängt in kom-

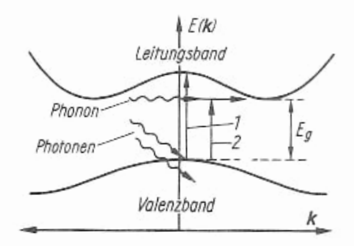


Abb. 8.8.3-2 Band-Band-Übergänge im $E(k)$-Diagramm; *1* direkter Übergang mit $\Delta E > E_g$; *2* indirekter Übergang mit $\Delta E = E_g$ unter Mitwirkung eines Phonons

plizierter Weise vom Zusammenspiel der aufgeführten Elementarprozesse ab. Primäre Fotoleitung liegt vor, wenn nur die unmittelbar durch das Licht erzeugten Ladungsträger in Richtung der Kontakte driften. Bei vielen Halbleitern tritt zusätzlich eine sekundäre Fotoleitung auf, durch die der primäre Fotostrom um mehr als das 10^4fache verstärkt werden kann. Dieser erwünschte Effekt kommt dadurch zustande, daß die an einer Elektrode abgeflossenen Ladungsträger am anderen Kontakt durch Injektion nachgeliefert werden. Das tritt vor allem dann ein, wenn die Rekombination dadurch behindert wird, daß eine Ladungsträgerart bevorzugt an Haftstellen eingefangen wird. Bei praktisch eingesetzten Fotoleitern, z. B. CdS, überwiegt gewöhnlich die Elektronenleitung, während die Löcher zum größten Teil an Haftstellen eingefangen werden. Die Verstärkung des Fotostroms infolge behinderter Rekombination erfolgt allerdings auf Kosten einer erhöhten Trägheit der Leitfähigkeitsänderung bei Belichtungswechsel.

Anwendungen findet die Fotoleitung in Widerstandszellen zur Messung von Lichtintensitäten. Die früher im sichtbaren Bereich eingesetzten Se- oder CdS-Fotowiderstände haben durch die Entwicklung von *Fotoelementen* (↑ 8.8.3.2.) an Bedeutung verloren. Weiterhin wichtig ist dagegen der Einsatz von Fotowiderständen als Infrarotdetektoren. Damit auch bei kleinen Quantenenergien eine Generation von Ladungsträgern möglich ist, muß man schmalbandige Halbleiter mit kleinem Bandabstand, z. B. PbS, InSb oder $Cd_xHg_{1-x}Te$, einsetzen. Zur Verminderung des Dunkelstroms infolge thermischer Anregung kühlt man die Fotowiderstände auf tiefe Temperaturen ab. Weiterhin werden fotoleitende Schichten in Fernsehaufnahmeröhren sowie in der Elektrofotografie (↑ 8.8.4.2.) eingesetzt.

8.8.3.2. Foto-EMK – Fotoelemente. Fotoelemente sind letztlich elektrische Generatoren, in denen Lichtenergie zur Erzeugung elektrischer Ströme genutzt wird. Notwendig hierzu ist der Aufbau von fotoelektromotorischen Kräften (Foto-EMK), auch Fotospannungen genannt. Sie bilden sich in Halbleitern oder Isolatoren aus, wenn die von den Lichtquanten erzeugten Ladungsträger beiderlei Vorzeichens unter dem Einfluß innerer Felder oder wegen ihrer unterschiedlichen Beweglichkeit getrennt werden.

8.8.3.2.1. *Fotoelemente mit pn-Übergang.* Fotoelemente auf der Basis von Halbleiterdioden mit *pn*-Übergang haben sowohl für meßtechnische Zwecke als auch für die Elektrizitätserzeugung aus der Energie der Sonnenstrahlung die größte Bedeutung erlangt. Das Prinzip eines Silicium-Fotoelements ist in Abb. 8.8.3-3 dargestellt. Wichtig für die Funktion ist, daß der *pn*-Übergang in relativ kleinem Abstand zur

Oberfläche des Halbleiters verläuft. Die in der oberen Si-Schicht durch das Licht erzeugten Ladungsträger werden dann im Feld des *pn*-Übergangs getrennt, indem die zum Übergang diffundierenden Minoritätsträger unter dem Einfluß des inneren Felds auf die andere Seite

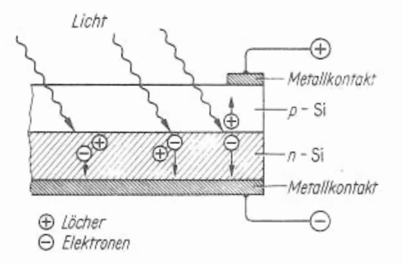


Abb. 8.8.3-3 Aufbau und Wirkungsweise eines Silicium-Fotoelements (schematisch)

gezogen werden (↑ 8.3.6.1.). Die Trennung der zusätzlich gebildeten Elektronen und Defektelektronen führt dazu, daß die Potentialbarriere am *pn*-Übergang teilweise abgebaut wird. Die Foto-EMK oder Fotospannung im Leerlauf ergibt sich aus der Differenz zwischen den Potentialbarrieren ohne bzw. bei Belichtung. Schließt man das Element über die metallischen Kontakte kurz, so fließt ein Strom, dessen maximale Dichte $j_{max} = eg_{ph}(L_p + L_n)$ beträgt. L_p und L_n sind die Diffusionslängen der Löcher bzw. Elektronen; sie sind gleich dem Produkt aus den Diffusionskoeffizienten (↑ 8.3.5.1.) und der mittleren Trägerlebensdauer (↑ 8.3.4.2.). Mit g_{ph} wird die Generationsrate von Ladungsträgern aufgrund der Photonenabsorption bezeichnet. Aus der Gleichung folgt, daß weitgehend perfekte, also einkristalline Halbleiterstrukturen notwendig sind, weil die Trägerlebensdauern mit dem Ansteigen von Gitterstörungen rasch absinken. Theoretisch ist mit Si-Fotoelementen bezogen auf irdische Sonnenstrahlung ein maximaler Wirkungsgrad von 22% erreichbar. Das liegt nahe dem Optimum, denn Halbleiter mit größerem Bandabstand würden zwar höhere Fotospannungen, jedoch nur eine kleinere Stromdichte liefern. Aus ökonomischen wie aus technologischen Erwägungen ist dem Silicium daher für diskrete Fotodioden für meßtechnische Zwecke ebenso wie für kleinere oder größere Solarbatterien der Vorzug zu geben. Selbst mit der ausgefeilten Si-Technologie wird der theoretische Wirkungsgrad jedoch nicht erreicht.

Als *Solarzellen* oder Sonnenbatterien werden großflächige Si-Fotoelemente, die passend in Reihe bzw. parallel geschaltet sind, bereits in großem Umfang eingesetzt. So sind die meisten Weltraumsatelliten mit ausklappbaren Flügeln ausgestattet, auf denen eine Vielzahl von Solarzellen angeordnet ist. Irdische Anwendungen in kleinerem Maßstab umfassen die Stromversorgung von Uhren sowie von meteorologischen Stationen oder Einrichtungen zur Nachrichtenübertragung in entlegenen, jedoch sonnenreichen Gegenden. An der Entwicklung von großflächigen Einheiten, die zur Energiegewinnung in sonnenreichen Gebieten auf der Erdoberfläche oder mit Hilfe großer Weltraumstationen geeignet sind, wird intensiv gearbeitet. Nachteilig sind die hohen Kosten. Man versucht daher alternativ, anstelle von einkristallinem Silicium Schichten aus amorphem, dotierbarem Si, das mit eingebautem Wasserstoff oder Fluor stabilisiert ist, für großflächige und billigere Solarzellen einzusetzen. Die bisher erreichten Wirkungsgrade liegen mit wenigen Prozent jedoch wesentlich niedriger.

Verwandt mit den Fotoelementen sind die *Fototransistoren*. Bei ihnen ist ein Teil der Oberfläche des Basisbereichs mit einem lichtdurchlässigen Fenster versehen. Bei Lichteinfall werden in der Basis Ladungsträgerpaare gebildet, die an den *pn*-Übergängen zum Emitter bzw. zum Kollektor getrennt werden. Dieser zusätzliche Basisstrom bedingt eine Änderung der Potentialdifferenzen an den Übergängen, insbesondere dem Emitter-Basis-Übergang, so daß der Stromfluß zwischen Emitter und Kollektor in empfindlicher Weise in Abhängigkeit von der Lichtintensität gesteuert wird. Fototransistoren ermöglichen eine optoelektronische Kopplung zwischen aktiven Bauelementen, d. h. einen Verzicht auf Eingangsnetzwerke, wenn sie in den Schaltungen mit Leuchtdioden (↑ 8.8.3.3.) eingesetzt werden.

8.8.3.2.2. *Sperrschichtfotoelemente.* Schon länger bekannt als die Fotoelemente mit *pn*-Übergang sind die Sperrschichtzellen, die im Prinzip aus einer *p*-leitenden Halbleiterschicht bestehen, die beidseitig dem gleichen mit Metall kontaktiert ist, wobei auf der einen Seite nur eine sehr dünne, lichtdurchlässige Schicht aufgebracht wird. Im stromlosen Zustand bilden sich *Sperrschichten* aus, indem Elektronen aus dem Metall in den Halbleiter übertreten (↑ 8.3.6.2.9.). Wird das Element nun durch die dünne Metallschicht belichtet, so wird die Symmetrie aufgehoben, da die gebildeten Ladungsträger getrennt werden, was einseitig zu einem Abbau der Potentialbarriere führt. Im Ergebnis baut sich eine Foto-EMK auf, und über die Kontakte kann ein Strom entnommen werden, welcher der Lichtintensität angenähert proportional ist. Analoge Elemente lassen sich auch mit *n*-leitenden Schichten herstellen. Sperrschichtfotoelemente wurden früher vielfach in Belichtungsmessern und anderen Meßgeräten eingesetzt. Bekannte Ausführungsformen benutzten Selen- oder Kupfer-(I)-oxidschichten. Neuerdings erlangen Sperrschichtelemente wieder Interesse,

da man hofft, durch Aufdampfen geeigneter Schichten auf Trägerfolien großflächige Solarbatterien herstellen zu können, die wesentlich billiger sind als einkristalline Elemente mit *pn*-Übergang. Die Wirkungsgrade liegen dafür allerdings erheblich niedriger. Für die Versorgung von Uhren oder Taschenrechnern mit Energie, wobei Tageslicht schon ausreicht, werden solche Zellen bereits eingesetzt.

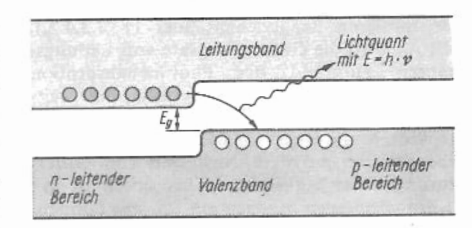
8.8.3.2.3. *Weitere Effekte.* Für die Halbleitermeßtechnik hat der *fotoelektromagnetische Effekt* (PME-Effekt) Bedeutung. Er tritt auf, wenn sich eine belichtete Halbleiterprobe in einem statischen Magnetfeld befindet. Man beobachtet transversal und longitudinal Fotospannungen, die ähnlich wie beim *Halleffekt* (↑ 8.3.5.2.) durch eine unterschiedliche Beeinflussung der diffundierenden Elektronen und Löcher im Magnetfeld zu erklären sind.

Beim *Dembereffekt* handelt es sich um das Auftreten einer kleinen Foto-EMK an belichteten homogenen Halbleiterproben. Die auch als *Kristallfotoeffekt* bekannte Erscheinung kommt durch Vorauseilen der Ladungsträgerart mit der größeren Beweglichkeit zustande. Bedeutung besitzt diese Erscheinung nur als Störungsquelle bei genauen Messungen elektrischer Halbleitereigenschaften.

8.8.3.3. Lumineszenz und Phosphoreszenz. Die Lumineszenz fester Stoffe besteht in einer Lichtemission bei Temperaturen weit unterhalb des Auftretens einer nennenswerten thermischen Quantenemission im betrachteten Frequenzgebiet. Bei festen Stoffen (es gibt auch Lumineszenz in Gasen oder Flüssigkeiten) kann die Lumineszenz entweder eine Kristall- oder eine Moleküleigenschaft sein, je nachdem ob die Elektronenübergänge, die zur Lichtemission führen, zwischen Elektronenzuständen der Bänder oder einzelner Molekülbausteine erfolgen. Lumineszenz kann auftreten, wenn genügend viele Elektronen in angeregten Niveaus und freie Plätze bzw. Löcher in energetisch tieferen Zuständen vorhanden sind. Anstelle der Lichtemission kann die beim Rückspringen der Elektronen freiwerdende Energie jedoch auch an Phononen (strahlungslose Übergänge) oder an andere Elektronen übertragen werden. Im letzten Fall können statt Lichtquanten Elektronen aus der Festkörperoberfläche austreten. Diese Erscheinung wird als *Augereffekt* bezeichnet. Die Analyse der Augerelektronen hat für die Untersuchung von Festkörperoberflächen große Bedeutung erlangt, da die Energieverteilung, die mit speziellen Augerspektrometern bestimmt wird, charakteristisch von der Zusammensetzung und Struktur der Festkörperoberfläche abhängt.

Die der Lumineszenz vorausgehende Elektronenanregung kann auf sehr verschiedene Weise erfolgen, beispielsweise durch Injektion von Ladungsträgern (Injektionslumineszenz), durch eingestrahltes Licht (Fotolumineszenz), durch Elektronenstrahlen (Katodolumineszenz), durch radioaktive Strahlung (Szintillationen) oder durch chemische Prozesse (Chemolumineszenz). Die praktischen Anwendungen der Lumineszenz reichen von der Optoelektronik bis zu den Leuchtstoffen von Lampen und den Bildschirmen der Fernseh- und Röntgentechnik.

8.8.3.3.1. *Leuchtdioden.* Läßt man durch eine Diode mit einem steilen *pn*-Übergang in Durchlaßrichtung Strom fließen, so kann eine intensive Emission von Lichtquanten auftreten (Abb. 8.8.3-4). Die Ladungsträger wandern dabei von beiden Seiten her aufeinander zu, und in der Nähe des Übergangs vereinigen sich Elektronen und Löcher durch Rekombination. Falls nicht strahlungslose Übergänge überwiegen, wird die freiwerdende Energie in Form von Licht ausgestrahlt. Die Energie der Quanten entspricht dabei entweder dem Bahnabstand, oder sie ist kleiner, wenn der Übergang zunächst von oder in Störniveaus der verbotenen Zone erfolgt. Um sichtbares Licht zu erhalten, muß man daher Halbleitermaterialien verwenden, deren Bandabstand mindestens 1,7 eV beträgt. Besonders geeignet sind *direkte Halbleiter* (↑ 8.2.1.4.3.), jedoch lassen sich auch mit *indirekten Halbleitern* brauchbare Lichtausbeuten erzielen, falls durch Einbau von Störstellen in genügender Konzentration dafür gesorgt wird, daß die Übergänge unter Beteiligung von Störniveaus erfolgen.



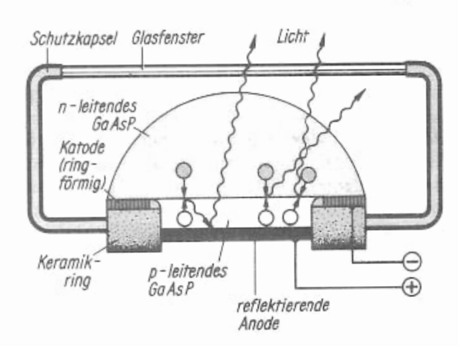


Abb. 8.8.3-4 Bandschema und Aufbau einer GaAsP-Fotodiode

Leuchtdioden bilden wichtige optoelektronische Bauelemente, die in Millionenstückzahlen gebaut und vorwiegend für die Signalanzeige in den verschiedensten Geräten und Anlagen eingesetzt werden. Für rotes Licht verwendet man InP oder GaAlP, für grünes Licht GaP oder GaAsP und gelb strahlen GaInP-Dioden. Lediglich die Entwicklung von blauen Strahlern bereitet Schwierigkeiten, obwohl Dioden aus dem leider schwer beherrschbaren SiC hierfür geeignet sein müßten.

Weitere Anwendungen der Leuchtdioden liegen im Gebiet schneller Schalter und in Verbindung mit Fototransistoren bei der Realisierung optisch gekoppelter elektronischer Schaltungen.

8.8.3.3.2. *Halbleiterlaser.* Laserdioden sind eine wichtige Weiterentwicklung der Leuchtdioden. Mittels sehr präziser Ausführung des *pn*-Übergangs, genauer Dimensionierung und Ausrichtung planparalleler Endflächen und ihrer Verspiegelung mit Mehrschichtsystemen erreicht man, daß sich in der Diode parallel zum Übergang ein Wellenfeld aufbaut (Abb. 8.8.3-5). Bei Stromfluß kommt es zur induzierten Emission kohärenter Lichtstrahlen. Aufgrund ihrer kleinen Abmessungen sind Halbleiterlaser als Lichtquellen für die Meßtechnik und für Videorecorder geeignet. Wegen ihrer schnellen Schaltzeit werden sie ferner für optoelektronische Schaltungen hoher Arbeitsgeschwindigkeit angewendet.

8.8.3.3.3. *Destriaueffekt (Leuchtkondensatoren).* Bettet man Zinksulfid, das mit Ag oder Cu dotiert ist, zwischen zwei leitfähige Kontakte ein, von denen einer lichtdurchlässig ist, so wird bei Anlegen einer Wechselspannung Licht emittiert. Die durchlässige Elektrode besteht gewöhnlich aus den transparenten und leitfähigen Stoffen SnO_2 und In_2O_3, die beispielsweise durch Aufstäuben abgeschieden werden. Neuerdings scheidet man sehr definierte ZnS-Schichten mit Hilfe von Molekularstrahlmethoden ab. Die so hergestellten Leuchtkondensatoren dürften gleichfalls für die Signalanzeige Interesse erlangen.

8.8.3.3.4. *Elektronenanregung in Ionenkristallen.* Die Kristalle der Alkalihalogenide sind normalerweise durchsichtig und farblos. Bestrahlt man sie jedoch mit genügend kurzwelligem UV-Licht oder mit Röntgen- bzw. Elektronenstrahlen oder erhitzt man sie in Alkalimetalldampf, so erlangen sie eine charakteristische Färbung. Beispielsweise färben sich LiCl gelb, NaCl gelbbraun, KCl rot und KBr blau. Pohl, Stasiw und Teltow führten das folgende Experiment durch, das als Grundversuch bekanntgeworden ist: Sie erhitzten einen KBr-Kristall zwecks Erhöhung der Leitfähigkeit auf 550 °C und legten über Metallelektroden eine Spannung von einigen 10^2 V an. Dabei wandert eine „blaue Wolke" in den Kristall ein, die sich bei Umpolung wieder zurückzieht. Durch zahlreiche Untersuchungen wurde seither sichergestellt, daß die Erscheinung durch Elektronen hervorgerufen wird, die an Gitterlücken, und zwar unbesetzten Plätzen des Anionenteilgitters eingefangen werden. Das ist verständlich, weil die umgebenden Kationen

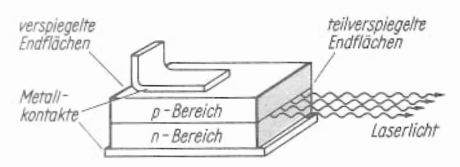


Abb. 8.8.3-5 Aufbau einer Laserdiode (schematisch)

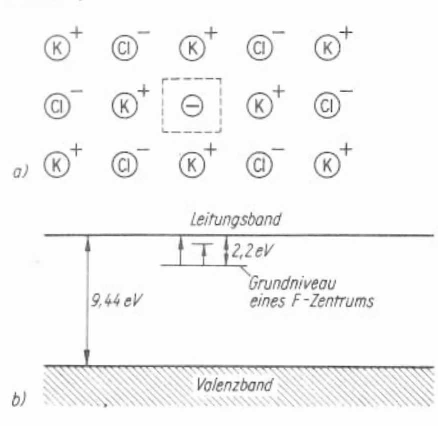


Abb. 8.8.3-6 a) F-Zentrum in Kaliumchlorid, b) Energiebandschema des Kaliumchlorids

eine Potentialmulde für das Elektron herstellen. Solche Punktdefekte werden als *F-Zentren* (Farbzentren) bezeichnet. Abb. 8.8.3-6 veranschaulicht die Lage des Grundniveaus und eines Anregungszustands für ein F-Zentrum in KCl. Inzwischen wurden auch weitere Farbzentren in Ionenkristallen identifiziert, so die F'-Zentren, die dem Einfang von zwei Elektronen an einer Anionenlücke entsprechen. Diese Elektronenniveaus sind bei vielen Erscheinungen der Lumineszenz und Phosphoreszenz in Festkörpern beteiligt.

8.8.3.3.5. *Leuchtstoffe und Kristallphosphore.* Die Phosphoreszenz bildet eine spezielle Form der Lumineszenz; sie ist durch ein Nachleuchten nach Aufhören der anregenden Prozesse über kürzere oder längere Zeit gekennzeichnet. Die Kristallphosphore zeichnen sich durch einen hohen Wirkungsgrad der Lichtemission aus. Kennzeichnend ist, daß die Lichtausbeute, das Nachleuchten und auch die Farbe des emittierten Lichts stark vom Einbau geeigneter Fremd-

atome oder *Aktivatoren* in das Gitter abhängen. Bekannte Kristallphosphore sind ZnS und CdS, die vorwiegend durch Cu oder Ag aktiviert werden. Die Aktivatoren wirken als Haftstellen, deren Abstand zu den Bandkanten bei lang nachleuchtenden Phosphoren groß gegen

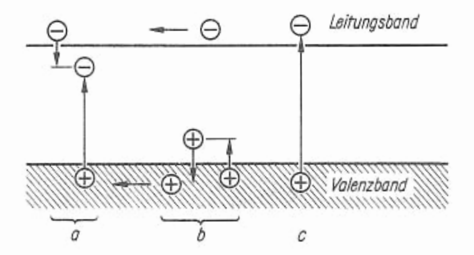


Abb. 8.8.3-7 Elementarprozesse bei der Lichtemission eines Sulfidphosphors; *a* Einfang von Elektronen und anschließend von Defektelektronen unter Lichtemission, *b* Einfang und Befreiung von Löchern an Haftstellen, *c* Anregung des Phosphors durch Lichtabsorption

die mittlere thermische Energie $k_B T$ bei Zimmertemperatur ist. Die thermische Befreiung der eingefangenen Träger wird dann stark verzögert, und ein langes Nachleuchten ist die Folge. Durch Erhitzen oder Einstrahlen von Licht läßt sich ein beschleunigtes Ausleuchten erzwingen. Bei vielen Kristallphosphoren sind mindestens zwei Arten von Störstellenniveaus am Zustandekommen der Lumineszenz beteiligt (Abb. 8.8.3-7). Kristallphosphore senden auch Lichtblitze aus, wenn Teilchen oder Quanten radioaktiver Strahlen auftreffen und ihre Energie ganz oder teilweise an das Gitter abgeben. In der Kernstrahlungsmeßtechnik wird dies zum Teilchen- oder Quantennachweis mit *Szintillationszählern* ausgenutzt (Abb. 8.8.3-8). Da die Sulfide das eigene Fluoreszenzlicht nur begrenzt durchlassen, setzt man zum Nachweis von γ-Quanten große Szintillatorkristalle aus NaI oder CsI ein, die mit Thallium aktiviert werden. Die relativ hohe Ordnungszahl der in diesen Stoffen enthaltenen Elemente ist erwünscht, weil daraus ein hohes Bremsvermögen für die Strahlen resultiert (↑ 3.3.3.). Wenn dies nicht so wichtig ist, dafür aber eine sehr kurze Impulszeit gefordert wird, zieht man fluoreszierende organische Molekülkristalle wie Stilben oder p-Terphenyl oder auch Lösungen von p-Terphenyl in Toluol bzw. in Plasten vor. Die Lichtblitze werden mit Hilfe eines Sekundärelektronenvervielfachers (SEV) in elektrische Impulse umgeformt und verstärkt. Ihre Intensität ist der abgegebenen Energie der Strahlenteilchen proportional, so daß eine Energiespektrometrie möglich ist. Neuerdings werden zum Strahlennachweis häufig *Halbleiterdetektoren* benutzt. Ihre Wirkungsweise beruht analog zur *Fotoleitung* (↑ 8.8.3.1.) auf einer impulsartigen Zunahme der Leitfähigkeit durch Generation von Ladungsträgern.

Die in Leuchtstofflampen, bei Röntgenschirmen und vor allem in den Fernsehbildröhren sowie in Elektronenmikroskopen und Oszillographen eingesetzten Leuchtstoffe sind prinzipiell den Kristallphosphoren gleichzusetzen. Bei diesen polykristallinen oder glasartigen Materialien konnten die Mechanismen der Anregung und Lumineszenz jedoch meist noch nicht in den Einzelheiten aufgeklärt werden, so daß ihre Herstellung und Verarbeitung noch vorwiegend nach empirischen Regeln erfolgt. Die Leuchtstoffe gehören zu recht verschiedenartigen Verbindungsgruppen. Wichtig sind Sulfide, Oxide, Silicate, die mit verschiedenartigen Metallen, darunter vor allem auch Vertretern der Seltenen Erden aktiviert werden. Durch passende Wahl des Grundmaterials und der Dotierung lassen sich die Farbe des emittierten Lichts sowie die Nachleuchtdauer in weiten Grenzen beeinflussen. Für Schwarz-Weiß-Bildröhren benutzt man zwei Stoffe mit Emission in den Komplementärfarben, während für Farbbildröhren wenigstens drei Stoffe in den entsprechenden Grundfarben erforderlich sind. Phosphore mit besonders großer Nachleuchtdauer auf der Basis von ZnS und BaS werden für die „Leuchtziffern" von Uhren und Armaturen eingesetzt. Eine zusätzliche Einlagerung von sehr geringen Mengen α-strahlender radioaktiver Stoffe hat wegen des Szintillationseffekts ein praktisches unbegrenztes Nachleuchten zur Folge.

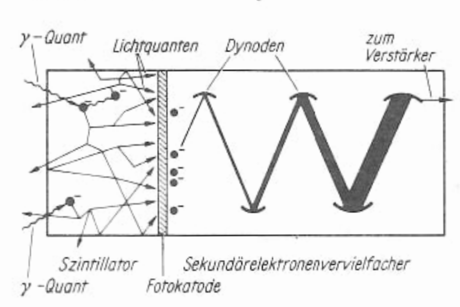


Abb. 8.8.3-8 Aufbau eines Szintillationszählers (schematisch)

8.8.4. Grundprozesse der Fotografie

Herausragende praktische Bedeutung haben schließlich auch die Festkörpereffekte, die bei fotografischen Prozessen eine Rolle spielen. Zu unterscheiden ist dabei zwischen der vorläufig noch dominierenden Fotografie mit lichtempfindlichen Silbersalzen und den neueren

Verfahren der Elektrofotografie, die für Kopierverfahren, aber zunehmend auch bei anderen Anwendungen eingesetzt werden.

8.8.4.1. Mechanismen der Fotografie mit Silbersalzen. Die lichtempfindlichen Schichten der Schwarzweißfotografie bestehen aus einer Emulsion von feinkörnigem Silberbromid AgBr oder gemischtem Silber-Bromid-Iodid. Bereits die Absorption von relativ wenigen Lichtquanten in einem Kristallkorn ruft Veränderungen hervor, die eine selektive Reduktion der belichteten, rund 10^4 Ag-Atome umfassenden Körner zu metallischem Silber bei Behandlung mit geeigneten Entwicklern auslösen.

Die vorwiegend von MOTT und GURNEY aufgestellte atomistische Theorie knüpft an die Erkenntnisse über Farbzentren an. Die Absorption eines Lichtquants regt ein Elektron in das Leitungsband an (Abb. 8.8.4-1a), das bei den isolierenden Silberhalogeniden normalerweise unbesetzt ist. Dieses Elektron wird an einer *Haftstelle*, wahrscheinlich meist an der Kornoberfläche, eingefangen (b). Eines der im AgBr stets in ausreichender Konzentration vorhandenen Silberionen auf einem *Zwischengitterplatz* wandert nun infolge der Anziehung durch das Elektron an die Haftstelle und vereinigt sich mit ihm zu einem neutralen Silberatom (c). Nach Absorption eines zweiten Lichtquants in dem gleichen Silberbromidkorn entsteht wiederum ein Elektron, das sich bevorzugt an das zuerst gebildete neutrale Ag-Atom anlagert und ein negatives Silberion Ag^- bildet (d). Dieses Ion zieht wiederum ein positiv geladenes Silberion von einem Zwischengitterplatz an und vereinigt sich (e) mit ihm zu einem Silberkeim Ag_2. Die Absorption weiterer Quanten führt zur Widerholung dieser Vorgänge, so daß schließlich bei genügend langer Belichtung sichtbare Silberkristalle entstehen. Schon nach der Absorption von nur 50 Quanten kann der belichtete Keim durch einen reduzierenden Entwickler vollständig in Silber umgewandelt werden. Dabei spielt insbesondere das *Keimwachstum* bei der Entstehung des latenten Bilds eine wichtige Rolle (↑ 6.4.2.).

8.8.4.2. Elektrofotografie. Als Elektrofotografie werden die Bildübertragungsverfahren zusammenfassend bezeichnet, bei denen bestimmte lichtelektrische Eigenschaften fester Stoffe genutzt werden. Man unterscheidet zwischen elektrostatischen Verfahren, Leitfähigkeitsverfahren und Methoden, die auf der Lichtempfindlichkeit der dielektrischen Polarisation beruhen. Betrachtet wird hier nur das elektrostatische Verfahren, das insbesondere in Verbindung mit der modernen Kopiertechnik einen außerordentlich breiten Einsatz gefunden hat. Abb. 8.8.4-2 zeigt schematisch die Verfahrensschritte. Die mit fotoleitendem Material, meist Selen oder Zinkoxid, überzogene Schichtoberfläche wird im Dunkeln elektrostatisch aufgeladen, was durch eine Koronaentladung geschehen kann (a). Im Verlauf der Belichtung wird die Oberflächenladung selektiv an den Stellen abgebaut, wo Fotoleitung auftritt (b). Hierdurch entsteht ein latentes Ladungsbild, das durch Bestreuen mit einem Farbpulver entwickelt wird, welches je nach dem Ladungszustand unterschiedlich haftet (c). Von der mehrfach benutzbaren Fotoleiterplatte wird das Pulverbild auf gleichfalls aufgeladenes Spezialpapier übertragen (d). Anschließend wird das Bild auf dem Papier thermisch oder chemisch fixiert (e). Zum Schluß reinigt man die Fotoleiterplatte von den Farbstoffresten (f). Weiterentwicklungen elektrofotografischer Verfahren ermöglichen auch die Wiedergabe von Farbbildern. Zweifellos bleibt noch viel zu tun, um die elektrofotografischen Methoden physikalisch aufzuklären und zu optimieren. Man darf erwarten, daß eine allmähliche Ablösung des knappen Silbers in vielen Teilgebieten der Fotografie möglich wird.

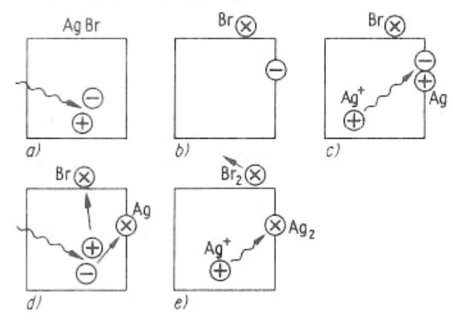


Abb. 8.8.4-1 Elementarprozesse bei der Entstehung des latenten fotografischen Bilds nach MOTT und GURNEY

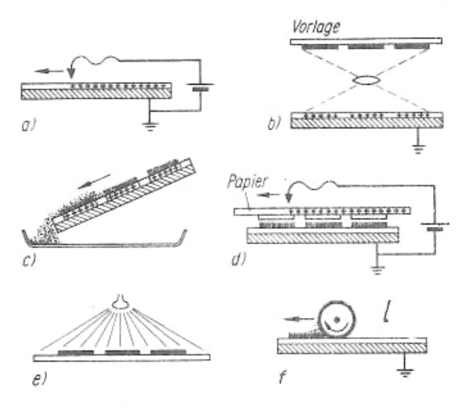


Abb. 8.8.4-2 Verfahrensschritte der Elektrofotografie im Aufladeverfahren nach SCHAFFERT

9. Materie unter extremen Bedingungen

9.1. Plasmazustand

9.1.1. Plasmazustand als vierter Aggregatzustand

9.1.1.1. Wesen und Vorkommen von Plasmen. So, wie durch ausreichende Energiezufuhr der feste in den flüssigen und dieser in den gasförmigen Aggregatzustand übergeführt werden kann, geht jedes Gas bei weiterer Temperaturerhöhung schließlich in den Plasmazustand über (Abb. 9.1.1-1). Während im festen, flüssigen und gasförmigen Aggregatzustand die *Elektronen* noch fest an die Atome oder Moleküle gebunden sind, bewegt sich im Plasmazustand ein Teil von ihnen frei und weitestgehend unabhängig von den zurückgebliebenen Atomrümpfen. Diese sind positiv geladen und werden als *Ionen* bezeichnet. Neben diesen positiven Ionen können sich durch Anlagerung von Elektronen an Neutralatome auch negative Ionen bilden. Durch das Auftreten freier Ladungen, von Elektronen und Ionen, und durch die damit verbundenen Besonderheiten unterscheidet sich der Plasmazustand grundsätzlich von den anderen Aggregatzuständen; man kann ihn deshalb als den *vierten Aggregatzustand* bezeichnen. Unter den Bedingungen unserer Umgebung an der Erdoberfläche werden Plasmen nur ausnahmsweise angetroffen; dagegen liegt im Kosmos, in den Fixsternen ebenso wie im interstellaren Raum der weitaus überwiegende Teil der stofflichen Materie im Plasmazustand vor (↑ 10.3.).

9.1.1.2. Entstehung des Plasmazustands. Um ein Gas in den Plasmazustand überzuführen, muß wenigstens ein Teil der Elektronen von den Atomen oder Molekülen abgetrennt werden. Dieser Vorgang wird als Ionisierung bzw. als *Ionisation* bezeichnet. Die für die Ablösung *eines* Elektrons notwendige Energie heißt *Ionisierungsenergie* (Tab. 9.1.1-2). Sie hängt eindeutig von der Stellung des zu ionisierenden Elements im Periodensystem der Elemente ab; z. B. haben die Alkalimetalle Li, Na, K, Rb, Cs in der äußeren Elektronenschale nur ein einziges, ziemlich locker gebundenes Elektron, das

Tab. 9.1.1-2 Ionisierungsenergie E_{in} des neutralen Atoms und des einfach geladenen Ions E_{i1} sowie niedrigste Anregungsenergie $E_{a\,min}$ des neutralen Atoms in eV für verschiedene Elemente

		E_{in}	E_{i1}	$E_{a\,min}$
Wasserstoff	H	13,5	–	10,2
	H_2	15,4		
Helium	He	24,5	54,2	19,8
Stickstoff	N	14,5	29,5	9,2
	N_2	15,8		
Sauerstoff	O	13,6	34,9	8,5
	O_2	12,5		
Neon	Ne	21,5	40,9	16,6
Natrium	Na	5,12	47,5	2,12
Argon	Ar	15,7	27,8	11,8
Kalium	K	4,32	31,7	1,55
Calcium	Ca	6,1	11,8	1,89
Eisen	Fe	7,83	16,5	
Kupfer	Cu	7,7	20,2	1,38
Krypton	Kr	14,0	26,4	9,9
Xenon	Xe	12,1	21,2	8,3
Caesium	Cs	3,9	23,4	1,48
Quecksilber	Hg	10,4	18,7	4,86

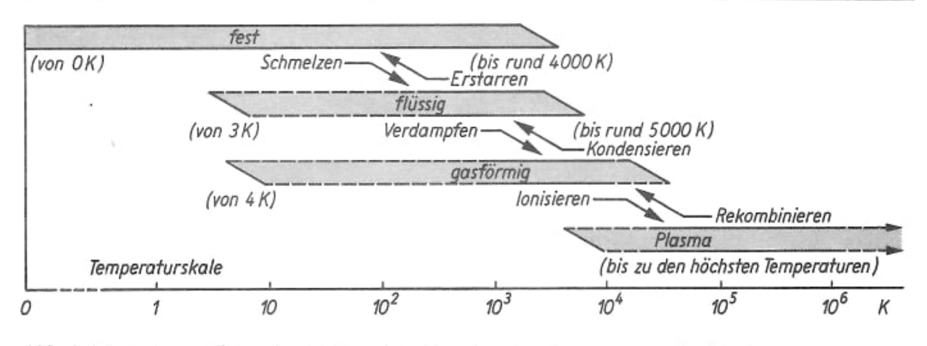


Abb. 9.1.1-1 Darstellung der Existenzbereiche der einzelnen Aggregatzustände

relativ leicht abgelöst werden kann († 4.3.3.). Ihre Ionisierungsenergien sind demzufolge niedrig. Die Edelgase He, Ne, Ar, Kr, Xe andererseits verfügen über abgeschlossene äußere Elektronenschalen. Die damit verbundene stabile Elektronenbindung bedingt vergleichsweise hohe Ionisierungsenergien. Innerhalb der einzelnen Gruppen des Periodensystems der Elemente nimmt die Ionisierungsenergie aufgrund der mit der Ordnungszahl wachsenden Abschirmung des Atomkernfelds durch die inneren Elektronen ab. Caesium als schwerstes Element der am leichtesten zu ionisierenden Alkalimetallgruppe hat daher von allen Elementen mit 3,9 eV die geringste Ionisierungsenergie. Bereits bei Temperaturen von 2000 bis 3000 K zeigt ein Caesiumgas typische Plasmaeigenschaften. Eine vollständige Ionisierung wird allerdings auch erst bei einigen 10^4 K erreicht. Daraus ersieht man, daß im Unterschied zu den Übergängen Festkörper → Flüssigkeit und Flüssigkeit → Gas die Grenze zwischen gasförmigem Aggregatzustand und Plasmazustand nicht scharf ist. Bei genügend hohen Temperaturen schließlich kann jeder beliebige Stoff nur im Plasmazustand existieren. Er besteht dann weitestgehend aus freien Elektronen und Ionen und enthält kaum noch ungeladene Teilchen, Atome oder Moleküle. Für Plasmen, die sich im thermischen Gleichgewicht befinden, kann der Anteil ionisierter Teilchen mit Hilfe der Boltzmannstatistik berechnet werden († 6.1.2.3.1.).

9.1.1.3. Verschiedene Ionisierungsmöglichkeiten. Um ein Gas zu ionisieren und damit in den Plasmazustand überzuführen, gibt es verschiedene Möglichkeiten. In Natur und Technik sind vor allem die Elektronenstoßionisierung und die Fotoionisierung weit verbreitet.

9.1.1.3.1. *Elektronenstoßionisierung.* Hierbei wird dem Atom die zur Abspaltung eines Elektrons erforderliche Energie bei einem Stoß mit einem schnellbewegten Elektron übertragen. Eine Ionisierung kann immer nur dann eintreten, wenn die beim Stoßvorgang auf das Atom übertragene Energie die Ionisierungsenergie übersteigt. Grundsätzlich ist auch eine Ionisierung des Atoms durch Ionen- oder Neutralteilchenstoß möglich. Allerdings ist der beim Stoß dieser schweren Stoßpartner auf ein Elektron übertragene Energieanteil verhältnismäßig gering. Er reicht gewöhnlich nicht zur Deckung der Ionisierungsenergie aus. Deshalb findet bei Ionen- oder Neutralteilchenstößen meist keine Ionisierung statt. Daraus wird die bevorzugte Rolle der Elektronenstoßionisierung im Plasma deutlich. Die dazu notwendigen Elektronen können außerhalb des Gases erzeugt und anschließend in das zu ionisierende Gas eingeschossen werden. Dieser Weg der *Ionisierung durch Elektronenstrahleinschuß* wird jedoch nur selten beschritten. Meist werden die im teilweise ionisierten Gas vorhandenen *Plasmaelektronen* zur Elektronenstoßionisierung herangezogen. Die zur Auslösung von Ionisierungsakten notwendige kinetische Energie wird ihnen dabei durch im Plasma selbst künstlich aufrechterhaltene elektrische Felder zugeführt. Da bei jedem Ionisierungsakt neue Elektronen frei werden und diese gleichfalls aus dem elektrischen Feld fortwährend Energie aufnehmen und damit ionisieren können, bildet sich sehr schnell eine regelrechte *Elektronenlawine* aus (Abb. 9.1.1-3). Bezeichnet n die Anzahl der Elektronen in der

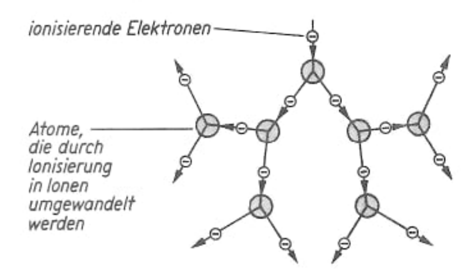


Abb. 9.1.1-3 Ausbildung einer Elektronenlawine

Volumeneinheit an einem Ort x im Plasma, so wird das Anwachsen der Elektronendichte in Richtung des elektrischen Felds, die mit der x-Richtung zusammenfallen soll, durch die Differentialgleichung $dn = \alpha n \, dx$ beschrieben. Der von der Stärke des elektrischen Felds, vom Gasdruck und von der Gasart abhängige Proportionalitätsfaktor α heißt *Elektronenstoßionisierungskoeffizient.* Er gibt an, wieviel Trägerpaare ein Elektron längs der Wegeinheit in Feldrichtung durch Ionisation erzeugt. Durch Integration dieser Differentialgleichung erhält man für die Zunahme der Elektronendichte n in der Lawine in Feldrichtung das Exponentialgesetz $n(x) \sim \exp(\alpha x)$.

Eine Sonderform der Elektronenstoßionisierung stellt die *thermische Ionisierung* dar. Bei ihr wird den Plasmaelektronen die zur Durchführung von Ionisationsakten erforderliche Energie durch einfache Erwärmung, d. h. *thermisch,* zugeführt. Die durch die Erwärmung immer heftiger werdende Bewegung *aller* im Gas enthaltenen Teilchen führt schließlich dazu, daß die bei Stößen von anfangs zufällig vorhandenen Elektronen mit Neutralteilchen übertragenen Energien zu deren Ionisierung ausreichen. Die neu entstandenen Elektronen gewinnen durch die Wärmezufuhr bald ebenfalls ausreichende kinetische Energie und vermögen weitere Ionisationsakte auszulösen. Bei extrem hohen Temperaturen über 10^6 K führen schließlich auch

Stöße zwischen den schwereren Teilchen, den Ionen und Neutralteilchen, zur Ionisierung.

9.1.1.3.2. *Fotoionisierung*. Hierbei wird dem Atom die zur Ionisation notwendige Energie mittels kurzwelliger Lichtstrahlung zugeführt. Da man nach der Quantentheorie die elektromagnetische Lichtstrahlung auch als eine Teilchenstrahlung aus *Lichtquanten* oder *Photonen* auffassen kann, bezeichnet man hier den Vorgang der Energieübertragung häufig auch als *Lichtquanten-* oder *Photonenstoß*. Wieder muß die bei einem solchen Stoß den Lichtquanten innewohnende Energie größer als die Ionisierungsenergie sein. Das ist gewöhnlich nur bei elektromagnetischer Strahlung relativ kurzer Wellenlänge der Fall. Daher kann eine Ionisierung nur mit ultraviolettem Licht oder mit Strahlung noch kürzerer Wellenlänge, also Röntgen- oder γ-Strahlen, erreicht werden.

9.1.1.4. Verschwinden des Plasmazustands. Im Plasma laufen auch Vorgänge ab, die der Ionisierung entgegenwirken und bei Abkühlung den Plasmazustand verschwinden lassen. In *Umkehrung* zur Ionisierung findet dabei eine *Wiedervereinigung* von Ionen und Elektronen zu neutralen Atomen statt, die als *Rekombination* bezeichnet wird. Bei zufälligen Zusammenstößen nicht zu energiereicher Elektronen und Ionen haften diese wegen ihrer ungleichnamigen Ladungen und der damit verbundenen elektrischen Anziehungskräfte leicht aneinander und bilden dann ein neutrales Gasatom. Dabei wird die zur vorangegangenen Ionisation aufgewendete Ionisierungsenergie wieder frei. Wird deshalb in einem vorhandenen Plasma die für die Ionisierung notwendige Energiezufuhr unterbrochen, so verschwinden die freien Elektronen und Ionen sehr bald wieder durch Rekombination. Bezeichnet n die Anzahl der Elektronen in der Volumeneinheit zur Zeit t im Plasma, so wird die Abnahme der Elektronendichte nach Aufhören der Energiezufuhr für die meisten Laborplasmen, in denen die Rekombination auf der Grundlage von Dreierstößen abläuft (↑ 9.1.2.1.2.), durch die Differentialgleichung $dn = \beta n^2 \, dt$ beschrieben. Der vom Gasdruck und von der Gasart abhängige Proportionalitätsfaktor β wird *Rekombinationskoeffizient* genannt. Die Integration dieser Differentialgleichung ergibt, daß die Elektronendichte n und damit auch die Ionendichte nach Einstellung der zur Ionisierung notwendigen Energiezufuhr gemäß dem Gesetz $n(t) = n_0/(1 + \beta n_0 t)$ vom Ausgangswert n_0 zeitlich abklingt.

Bei extrem hohen Temperaturen tritt die Rekombination in den Hintergrund, da die Elektronen und Ionen bei zufälligen Zusammenstößen infolge ihrer hohen kinetischen Energie nicht mehr aneinander haften bleiben. Daher wird bei solchen Temperaturen eine vollständige Ionisierung möglich. Im Plasma sind dann nur noch geladene Teilchen, Ionen und Elektronen vorhanden. Da die chemischen Bindungsenergien geringer sind als die Ionisierungsenergien, gibt es dann auch keine Moleküle mehr, und das Plasma enthält neben den Elektronen nur positive Atomionen, die auch mehrfach geladen sein können.

9.1.2. Elementarprozesse und Strahlung im Plasma

9.1.2.1. Wichtige Elementarprozesse. Ein Gas im Plasmazustand besteht aus einem Gemisch von Elektronen, Ionen und noch nicht ionisierten Atomen oder Molekülen des Grundgases. Zwischen allen diesen Teilchen finden fortwährend Zusammenstöße statt. Da sich die schwereren Teilchen, die Ionen, Atome oder Moleküle, außerdem in verschiedenen angeregten Energiezuständen befinden können und damit Strahlung emittieren oder absorbieren können, spielt auch die Wechselwirkung zwischen diesen Teilchen und der im Plasma vorhandenen Strahlung eine Rolle. Die verschiedenen Prozesse, die bei Teilchenstößen oder bei der Wechselwirkung mit der Strahlung vor sich gehen, werden als *Elementarprozesse* bezeichnet (Abb. 9.1.2-1).

Am häufigsten sind solche Stöße zwischen den Teilchen des Plasmas, bei denen nur sehr wenig Energie ausgetauscht wird. Diese elastischen Stöße bestimmen das kinetische Verhalten der Teilchen sowohl im gasförmigen als auch im Plasmazustand. Da sie wegen des geringen mit ihnen verbundenen Energieaustauschs nicht zur Bildung von geladenen Teilchen (Ionisierung) führen können, sind sie für den Plasmazustand nicht allzu bedeutungsvoll. Große Bedeutung für den Plasmazustand haben indessen die Stöße, die vermöge eines größeren Energieaustauschs zu tiefergreifenden Veränderungen an den Teilchen führen können. Sie werden deshalb als *inelastische Stöße* bezeichnet. Zu ihnen gehören unter anderem alle Stöße, die zur Elektronenanregung, zur Ionisation oder zur Rekombination führen.

9.1.2.1.1. *Ionisation*. Dieser Elementarprozeß wird vorzugsweise durch *Elektronen-* oder *Photonenstoß* hervorgerufen (↑ 9.1.1.3.). Dabei muß auf den Stoßpartner, das Atom oder das Molekül, wenigstens so viel Energie übertragen werden, wie zur Ablösung eines Elektrons nötig ist. Kann diese Ionisierungsenergie nicht aufgebracht werden, so findet keine Ionisierung statt; der Stoß bleibt dann elastisch. Wird beim Stoß auf das Atom eine größere Energie als die Ionisierungsenergie übertragen, so kommt der Überschuß vor allem dem abgetrennten Elektron als kinetische Energie zugute.

Ähnlich wie neutrale Atome oder Moleküle können auch bereits vorhandene Ionen ionisiert werden. Durch die Abtrennung eines oder mehrerer weiterer Elektronen entstehen dann *doppelt* oder *mehrfach geladene Ionen.* Man bezeichnet solche Vorgänge als *Mehrfachionisierung.* Allerdings sind zur Abtrennung weiterer Elektronen beträchtlich größere Ionisierungsenergien nötig, so daß solche höheren *Ionisationsstufen* erst bei höheren Temperaturen auftreten. Das Plasma enthält dann außer Neutralteilchen, Elektronen und einfach geladenen Ionen auch mehrfach geladene Ionen. Man sieht hieraus, daß die Zusammensetzung des Plasmas stark von seiner Temperatur abhängt. Sie verändert sich mit wachsender Temperatur zugunsten immer höherer Ionisationsstufen (Abb. 9.1.2-2).

Abb. 9.1.2-1 Die wichtigsten Elementarprozesse im Plasma und deren Umkehrungen: a_1) Anregung durch Elektronen- oder Photonenstoß, a_2) Rückkehr in den Grundzustand, b_1) Ionisation durch Elektronen- oder Photonenstoß, b_2) Rekombination durch Dreier- oder Zweierstoß, c_1) Umladung, c_2) Anlagerung, d_1), d_2) Ionisierung und nachfolgende Dissoziation von Molekülen; *E* Elektron, *A* Atom, *I* Ion, *Ph* Photon, *M* Molekül, (*s*) schnell, (*l*) langsam, (*G*) Grundzustand, (*a*) angeregt, (*i*) ionisiert, (*b*) bewegt

9.1.2.1.2. *Rekombination.* Dieser Elementarprozeß ist der *Umkehrprozeß* der Ionisation. Er besteht in der *Wiedervereinigung* zweier entgegengesetzt geladener Teilchen, von Ion und Elektron, zu einem neutralen Atom. Eine Rekombination ist nur möglich, wenn das entstehende neutrale Atom die beim Rekombinationsvorgang freiwerdende Energie abführen kann. Diese ist gewöhnlich etwas größer als die Ionisierungsenergie, da das rekombinierende Elektron zusätzlich kinetische Energie, die es im Plasma gewonnen hat, mitbringt. Die genannte Bedingung führt zu zwei verschiedenen Mechanismen der Rekombination, die als Strahlungsrekombination und Dreierstoßrekombination bezeichnet werden. Bei der *Strahlungsrekombination,* bei der ein Elektron mit einem Ion zusammenstößt und die man daher auch als *Zweierstoßrekombination* bezeichnet, wird die freiwerdende Energie durch Emission eines Lichtquants abgestrahlt. Sie ist damit der direkte Umkehrprozeß der Fotoionisierung. Bei der *Dreierstoßrekombination* hingegen, die nur dann eintritt, wenn gleichzeitig ein Ion und *zwei* Elektronen zusammenstoßen, wird die freiwerdende Energie von einem der beiden Elektronen aufgenommen. Die Funktion dieses Elektrons kann auch ein beliebiger anderer Körper übernehmen, z. B. die Wand des Gefäßes, in dem das Plasma eingeschlossen ist. Man spricht dann im Unterschied zu der im Inneren des Plasmas stattfindenden *Volumenrekombination* von einer *Wand-* oder *Oberflächenrekombination.* Der Vorgang der Dreierstoßrekombination ist der direkte Umkehrprozeß zur Elektronenstoßionisierung.

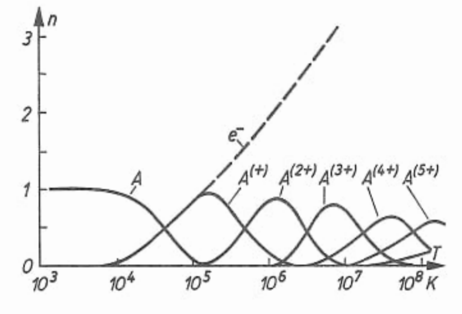


Abb. 9.1.2-2 Zusammensetzung des Plasmas in Abhängigkeit von der Temperatur (qualitativ); e^- Elektronen, *A* neutrale Atome, $A^{(+)}$ 1fach, $A^{(2+)}$ 2fach usw. positiv geladene Ionen

9.1.2.1.3. *Anregung*. Bei diesem Elementarprozeß handelt es sich ebenfalls um einen inelastischen Stoß. Im Gegensatz zur Ionisation sind die hierbei übertragenen Energien jedoch kleiner als die Ionisierungsenergie, so daß keine Elektronen von den Atomen oder Molekülen abgelöst werden können. Vielmehr werden durch die anregenden Stöße die Atome oder Moleküle in höhere Energiezustände übergeführt, von denen sie nach kurzer Zeit spontan unter Aussendung einer Lichtstrahlung wieder in ihren Grundzustand zurückkehren. Die zur Anregung der verschiedenen höheren Energiezustände der Atome oder Moleküle erforderlichen diskreten Energien heißen *Anregungsenergien* (↑ Tab. 9.1.1-2). Da bei der Anregung keine geladenen Teilchen erzeugt werden, ist sie für den Plasmazustand untypisch. Vielmehr sind Anregungsvorgänge schon im gasförmigen Aggregatzustand möglich. Da aber die Anregungsenergien eines Gases immer kleiner als die Ionisierungsenergie sind, finden in einem Plasma zwangsläufig auch Anregungsvorgänge mit großer Häufigkeit statt. Sie beruhen ebenso wie die Ionisierung vorzugsweise auf Elektronen- oder Lichtquantenstoß und sind die Ursache für einen Teil der Lichtstrahlung des Plasmas. Diese besteht aus *diskreten Spektrallinien* und wird daher als *Linienstrahlung* bezeichnet (↑ 4.3.1. und 9.1.2.2.).

Schließlich ist bei ausreichender Energie der stoßenden Elektronen oder Lichtquanten auch eine Anregung der im Plasma vorhandenen Ionen möglich; dabei werden die bedeutend kurzwelligeren Spektrallinien ionisierter Atome emittiert.

9.1.2.1.4. *Weitere Elementarprozesse*. Bei genügender Annäherung von geladenen Ionen an neutrale Atome kann ein Ladungsaustausch zwischen beiden Teilchen eintreten. Dieser Vorgang wird *Umladung* genannt. Ein anderer Elementarprozeß ist die *Dissoziation*, die in der Regel bei der Ionisierung von mehratomigen Molekülen auftritt (↑ Abb. 9.1.2-1). Dabei zerfällt das ionisierte Molekül oder Molekülion sehr bald in ein Ion und ein oder mehrere ungeladene Neutralteilchen. Auch eine *Anlagerung* von Elektronen an neutrale Atome oder Moleküle unter Bildung negativer Ionen wird häufig beobachtet. Sie findet jedoch nur in solchen Gasen statt, deren Bausteine über eine besonders große *Elektronenaffinität* verfügen, z. B. in Sauerstoff.

9.1.2.2. Strahlung im Plasma. Die vom Plasma erzeugte und ausgesandte Strahlung ist elektromagnetischer Natur und über weite Gebiete des Spektrums verteilt. Die Plasmastrahlung läßt sich am besten nach den zugrundeliegenden Erzeugungsmechanismen klassifizieren. Danach unterscheidet man zwischen Linienstrahlung, Rekombinationsstrahlung und Bremsstrahlung.

9.1.2.2.1. Bei der *Linienstrahlung* handelt es sich um *diskrete Spektrallinien*, die nach vorangegangener Anregung von Atomen, Molekülen oder Ionen (↑ 9.1.2.1.3.) ausgesendet werden. Die von den angeregten Teilchen bei der Rückkehr in den Grundzustand emittierten Spektrallinien bilden regelrechte Liniensysteme, die als *Atom-*, *Molekül-* bzw. *Ionenspektren* bezeichnet werden und für die betreffenden Atom-, Molekül- und Ionenarten spezifisch sind (↑ 4.3.3.). Während die Atomspektren im wesentlichen im sichtbaren und ultravioletten Spektralbereich liegen, erstrecken sich die Ionenspektren über den ultravioletten Spektralbereich bis in den Bereich der weichen Röntgenstrahlung. Molekülspektren sind ausschließlich im infraroten Spektralbereich zu finden und bestehen aus Folgen außerordentlich vieler und sehr dicht beieinanderliegender Spektrallinien, den *Molekülbanden*. In der verschiedenen Intensität dieser Spektren spiegelt sich die stark temperaturabhängige Zusammensetzung des Plasmas aus Molekülen, Atomen und einfach oder mehrfach geladenen Ionen wider.

9.1.2.2.2. *Rekombinationsstrahlung*. Sie entsteht bei der Zweierstoßrekombination von Elektronen und Ionen, die daher auch als *Strahlungsrekombination* bezeichnet wird (↑ 9.1.2.2.2.). Dabei wird die beim Rekombinationsvorgang freiwerdende Energie abgestrahlt. Weil die rekombinierenden Elektronen des Plasmas infolge ihres vorangegangenen Aufenthalts im Plasma über unterschiedliche kinetische Energien verfügen, kann sich keine diskrete Linienstrahlung ausbilden. Vielmehr entsteht eine kontinuierliche Strahlung, die als *Rekombinationskontinuum* bezeichnet wird. Dieses hat eine langwellige Grenze, die durch die Ionisierungsenergie des bei der Rekombination entstehenden Atoms festgelegt ist.

9.1.2.2.3. *Bremsstrahlung*. Sie entsteht, weil jede verzögerte Bewegung geladener Teilchen mit einer Ausstrahlung elektromagnetischer Strahlung verbunden ist, deren durch die Energie pro Zeiteinheit und pro Flächeneinheit auf der Einheitskugel gegebene Intensität $I = 2e^2\ddot{r}^2/(4\pi\varepsilon_0\, 4\pi\, 3c^2)$ beträgt, wenn e die Elementarladung, c die Vakuumlichtgeschwindigkeit und $\ddot{r}$ die Verzögerung des geladenen Teilchens bedeuten. Bei der Wechselwirkung von Elektronen und Ionen im Plasma finden solche Verzögerungsprozesse laufend statt, so daß es im Plasma zur Erzeugung einer intensiven *Bremsstrahlung* kommt. Da auch die Wellenlänge der so erzeugten elektromagnetischen Strahlung von der Stärke der jeweiligen Verzögerung abhängt, tritt die Bremsstrahlung zwangsläufig als kontinuierliche Strahlung in Erscheinung und erstreckt sich über weite Gebiete des Spektrums.

9.1.2.2.4. *Aufbau des Spektrums der Plasma-*

strahlung, Emissions- und Absorptionsverhalten. Gewöhnlich treten in Plasmen Linien-, Rekombinations- und Bremsstrahlung gleichzeitig auf, wobei ihr jeweiliger Anteil stark von der Temperatur und von einigen weiteren Bedingungen abhängt. Dadurch weist das Spektrum der Plasmastrahlung im Vergleich zum Spektrum von Festkörpern, Flüssigkeiten und Gasen zumeist eine ziemlich komplexe Struktur auf (Abb. 9.1.2-3). Mit steigender Temperatur tritt die Linienstrahlung allmählich in den Hintergrund, Rekombinations- und Bremsstrahlung dominieren immer mehr. Bei extrem hohen Temperaturen schließlich verschwindet die Linienstrahlung gänzlich, da dann alle Atome vollständig ionisiert sind, das Plasma nur noch aus freien Elektronen und nackten Atomkernen besteht und damit keine Anregungen mehr stattfinden können. Bei noch höheren Temperaturen tritt auch die Rekombinationsstrahlung in den Hintergrund, da wegen der hohen kinetischen Energie der Elektronen kaum noch Rekombinationsakte stattfinden können. Die Strahlung des Plasmas besteht dann fast ausschließlich aus Bremsstrahlung.

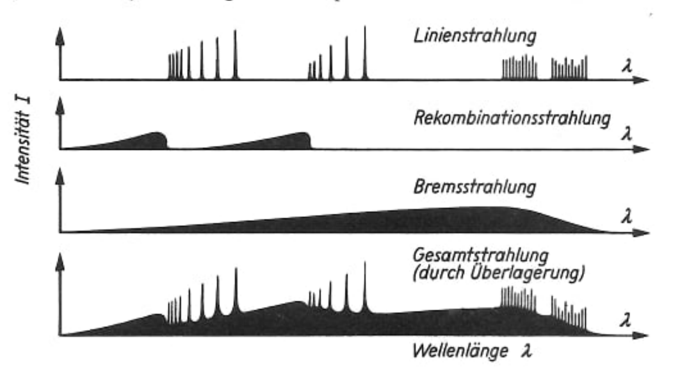


Abb. 9.1.2-3 Aufbau des Spektrums der Plasmastrahlung

Das Plasma kann elektromagnetische Strahlung auch absorbieren und wird dann als *optisch dicht* bezeichnet; andernfalls ist es *optisch dünn.* Eine starke Absorption ist allerdings nur für solche Wellenlängen möglich, die vom Plasma auch emittiert werden können, weil die zur Strahlungsemission führenden Vorgänge nur die Umkehrprozesse zu den der Strahlungsabsorption zugrunde liegenden Vorgängen sind. Außer von der Wellenlänge hängt die Strahlungsabsorption im Plasma noch von dessen Schichtdicke ab.

Die im Inneren eines optisch dichten Plasmas erzeugte Strahlung kann wegen der starken Absorption nicht ohne weiteres austreten, weil die durch Emission entstandene Strahlung vom Plasma selbst wieder absorbiert wird; sie kann sogar mehrere Male emittiert und wieder absorbiert werden. Da zwischen jedem Absorptions- und dem nachfolgenden Emissionsakt eine gewisse Zeit vergeht, braucht die Strahlung bei größerer Ausdehnung des Plasmas relativ lange Zeit, um aus dem Inneren an die Oberfläche und damit nach außen zu gelangen. Man bezeichnet das schrittweise Nachaußenwandern der Strahlung vermittels einer Vielzahl aufeinanderfolgender Absorptions- und Emissionsakte als *Strahlungsdiffusion.* Diese erfolgt sehr viel langsamer als die Ausbreitung der elektromagnetischen Strahlung im freien Raum, die bekanntlich mit Lichtgeschwindigkeit vor sich geht.

9.1.3. Bemerkenswerte Eigenschaften des Plasmazustands

9.1.3.1. Thermodynamisches Verhalten des Plasmas. Genauso wie im gasförmigen Aggregatzustand, befinden sich auch im Plasmazustand die Stoffteilchen, d. h. neben Atomen und Molekülen auch Elektronen, Ionen u. a., in ständiger ungeordneter *thermischer Bewegung*; nur daß diese Wärmebewegung der Teilchen im Plasmazustand wegen seiner höheren Temperaturen bedeutend heftiger als im gasförmigen Zustand erfolgt. Die Geschwindigkeiten der zahlreichen Plasmateilchen sind dabei nicht von einheitlicher Größe, sondern um einen Mittelwert, um die *mittlere thermische Geschwindigkeit*, verteilt. Die Funktionen, die die Verteilung der Geschwindigkeiten der einzelnen Teilchensorten beschreiben, werden *Verteilungsfunktionen* genannt. Sie enthalten als wichtigsten Parameter die Temperatur.

Im gasförmigen Aggregatzustand sind die Verhältnisse recht einfach. Dort bildet sich infolge der zahlreichen Stöße zwischen den gleichartigen Teilchen sehr bald eine stationäre, d. h. zeitlich nicht mehr veränderliche Geschwindigkeitsverteilung heraus. Sie wird nach

ihrem Entdecker als *Maxwellverteilung* bezeichnet (Abb. 9.1.3-1; ↑ 6.2.4.). Bedeuten n die Gesamtzahl der fraglichen Teilchen in der Volumeneinheit und dn die Anzahl von Teilchen in der Volumeneinheit, deren Geschwindigkeit v im Intervall zwischen v und $v + dv$ liegt, so

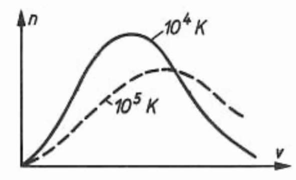


Abb. 9.1.3-1 Im isothermen Plasma meist verwirklichte Maxwellverteilung der Teilchengeschwindigkeiten; Teilchenzahl n und Geschwindigkeit v in relativen Einheiten

gilt das Maxwellsche Geschwindigkeitsverteilungsgesetz oder die Geschwindigkeitsverteilungsfunktion dn/dv:

$$\frac{dn}{dv} = \frac{4nv^2}{\sqrt{\pi}} \left(\frac{m}{2k_BT}\right)^{3/2} e^{-mv^2/2k_BT},$$

$$\bar{v} = \sqrt{\frac{8k_BT}{\pi m}}, \qquad v_n = \sqrt{\frac{2k_BT}{m}}.$$

Es bedeuten: $\bar{v}$ die mittlere thermische Geschwindigkeit, v_n die wahrscheinlichste Geschwindigkeit, n die Anzahl der Teilchen, k_B die Boltzmannkonstante, m die Teilchenmasse und T die absolute Temperatur. Aus einer solchen Maxwellverteilung lassen sich die mittlere thermische Geschwindigkeit $\bar{v}$ der Teilchen und die wahrscheinlichste Geschwindigkeit v_n bestimmen, die dem Maximum der Verteilungsfunktion entspricht.

Im Plasmazustand hingegen liegen die Verhältnisse sehr viel komplizierter, da sich in ihm oftmals gar keine stationären Gleichgewichtszustände ausbilden können. Der auf stationäres Verhalten zugeschnittene Begriff der Temperatur wird dadurch für das Plasma oft gegenstandslos.

Am einfachsten liegen noch die Verhältnisse bei *sehr dichten und verhältnismäßig kalten Plasmen.* Wegen der zahlreichen energieaustauschenden Stöße in ihnen stellt sich in relativ kurzer Zeit ein stationäres Gleichgewicht zwischen *allen* Plasmateilchen ein. Man nennt einen solchen Gleichgewichtszustand ein *vollständiges thermodynamisches Gleichgewicht*. Bei ihm ist man berechtigt, von einer *Temperatur des Plasmas* zu sprechen, das dann *isothermisch* genannt wird.

Bei *weniger dichten und dafür heißeren Plasmen* hingegen kann sich oft nur ein stationäres Gleichgewicht innerhalb einzelner Teilchensorten, der Atome, der Elektronen oder der Ionen, ausbilden, nicht aber zwischen den Teilchensorten. In diesem Fall spricht man von einem *partiellen thermodynamischen Gleichgewicht*, wobei streng zwischen *Neutralteilchentemperatur*, *Ionentemperatur* und *Elektronentemperatur* unterschieden werden muß (Abb. 9.1.3-2); z. B. kann der Unterschied zwischen Ionen- und Elektronentemperatur in solchen Plasmen einige 10^4 bis 10^5 K betragen. Plasmen dieser Art werden als *nichtisothermisch* bezeichnet.

In *sehr dünnen und heißen Plasmen* schließlich ist eine stationäre Gleichgewichtseinstellung auch innerhalb der einzelnen Teilchensorten in annehmbaren Zeiten nicht mehr möglich. Der Begriff der Temperatur wird dann völlig gegenstandslos und darf überhaupt nicht mehr verwendet werden. Eine thermodynamische Beschreibung eines solchen Plasmas ist unmöglich.

Ob sich in einem Plasma zwischen den Teilchen ein thermodynamisches Gleichgewicht ausbilden kann oder nicht, hängt außer von seiner Dichte und der damit zusammenhängenden Häufigkeit energieaustauschender Stöße auch noch von der Zeit ab, die für den Energieaustausch zur Verfügung steht; z. B. befinden sich Plasmen, die nur sehr kurze Zeit von der Größenordnung Mikrosekunden existieren, praktisch nie in einem thermodynamischen Gleichgewicht. Andererseits können sehr langlebige und sich nur außerordentlich langsam verändernde Plasmen auch bei geringer Teilchendichte als isothermisch angesehen werden. In solchen Plasmen finden während ihrer langen Lebenszeit trotz geringer Stoßhäufigkeit genügend viele energieaustauschende Stöße statt, um die Ausbildung eines vollständigen thermodynamischen Gleichgewichts zu ermöglichen.

Nur wenn sich ein Plasma im vollständigen thermodynamischen Gleichgewicht befindet, wenn es isothermisch ist, kann es mit thermodynamischen Vorstellungen und Begriffen, wie Temperatur oder spezifische Wärme, beschrie-

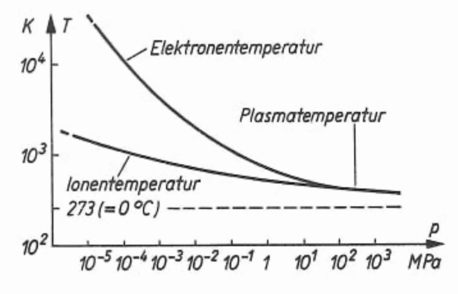


Abb. 9.1.3-2 Verlauf von Elektronen-, Ionen- und Plasmatemperatur in Abhängigkeit vom Druck (qualitativ)

ben werden. Aber auch hierbei ist wegen der komplizierten Vorgänge im Plasma äußerste Vorsicht am Platz; z. B. zeigt die spezifische Wärme eines isothermen Plasmas in Abhängigkeit von der Temperatur viele ausgeprägte Maxima und Minima (Abb. 9.1.3-3). Dies läßt sich folgendermaßen verstehen: Unter *spezifischer Wärmekapazität* versteht man die Wärmemenge, die der Masseeinheit eines Stoffs zugeführt werden muß, um seine Temperatur um 1 K zu erhöhen. Dabei muß anteilig für jedes Teilchen des Stoffs im Mittel eine gewisse kleine Energie je Grad Temperaturerhöhung aufgewendet werden. Da nun bei höheren Temperaturen in merklichem Umfang energieverzehrende Ionisationsprozesse stattfinden und dabei zusätzlich die Anzahl der Plasmateilchen anwächst, weil ein Atom bei der Ionisation zwei Teilchen, ein Ion und ein Elektron, ergibt, muß bei diesen Temperaturen zur weiteren Temperaturerhöhung eine größere Energie als gewöhnlich zugeführt werden. Das drückt sich im Verlauf der spezifischen Wärmekapazität als Anstieg aus. Bei weiterer Temperaturerhöhung sinkt die spezifische Wärmekapazität ab, um bei noch höheren Temperaturen, bei denen bereits Mehrfachionisierungen auftreten können, erneut steil anzusteigen. Deshalb zeigt der Verlauf der spezifischen Wärme eine Folge von Maxima, die in richtiger Reihenfolge das Auftreten einfacher, doppelter, dreifacher Ionisierung usw. ausweisen.

9.1.3.2. Quasineutralität und Mikrofeld. Bei der Ionisation werden die Ladungsträger, die ladungstragenden Elektronen und Ionen, immer paarweise gebildet, so daß das Plasma als Ganzes *elektrisch neutral* bleibt. Da sich jedoch im Plasma die Ladungsträger weitestgehend frei bewegen können, überwiegen oft innerhalb kleiner Volumenbereiche im Plasma Teilchen der einen oder anderen Ladungsart. Diese Volumenelemente sind dann nicht mehr elektrisch neutral, sondern enthalten Überschußladungen eines Vorzeichens. Zwischen Volumenelementen mit Überschußladungen verschiedenen Vorzeichens bilden sich sehr starke elektrische Felder aus, die eine weitere *Ladungstrennung* verhindern und damit bewirken, daß die im großen vorhandene Neutralität des Plasmas auch im kleinen nicht wesentlich gestört wird. Diese Eigenschaft des Plasmas, in jedem Volumenelement nahezu die gleiche Anzahl positiver und negativer Ladungsträger zu enthalten, bezeichnet man als *Quasineutralität* (↑ 8.3.3.4.). Sie ist eine der wichtigsten und bemerkenswertesten Eigenschaften des Plasmazustands und kann, wenn n_- die Anzahl der negativen und n_+ die der positiven Teilchen je Volumeneinheit bedeuten, durch $n_- \approx n_+$ oder $|n_- - n_+| \approx 0$ ausgedrückt werden.

Unabhängig davon existieren natürlich zwischen den geladenen Teilchen elektrische Felder. Durch Überlagerung der Felder aller im Plasma chaotisch verteilten geladenen Teilchen entsteht schließlich ein elektrisches Feld, dessen Stärke und Richtung von Ort zu Ort variiert.

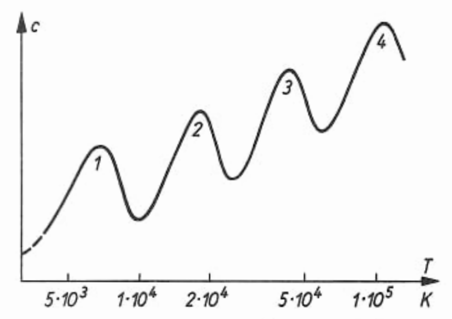


Abb. 9.1.3-3 Temperaturabhängigkeit der spezifischen Wärme c (in relativen Einheiten) eines isothermen Plasmas (qualitativ)

Es wird als *Mikrofeld* bezeichnet. Wegen der unregelmäßigen thermischen Bewegung der Ladungsträger ändert sich das Mikrofeld auch an ein und demselben Ort in schneller zeitlicher Folge. Die Stärke des Mikrofelds hängt nach statistischen Berechnungen allein von der Anzahl $n_- \approx n_+$ der Ladungsträger in der Volumeneinheit ab; für den Mittelwert des Absolutbetrags der Stärke des elektrischen Mikrofelds gilt $|E| = s\,e\,n_-^{3/2}$, wenn e die Elementarladung, n_- die Anzahl der negativen Teilchen je Volumeneinheit und s einen Zahlenfaktor zwischen 10 und 20 bedeuten, dessen genauer Wert von der der Berechnung zugrunde gelegten Ladungskonfiguration abhängt. Für typische Laborplasmen lassen sich so Werte zwischen 10^0 und 10^2 Vm^{-1} ausrechnen.

9.1.3.3. Leitfähigkeit, Stromtransport, Polarisation. Im Unterschied zum gasförmigen Aggregatzustand hat der Plasmazustand wegen der nahezu völlig frei beweglichen Ladungsträger eine bemerkenswerte *elektrische Leitfähigkeit.* Daher kann durch ein Plasma elektrischer Strom fließen. Grundsätzlich stellt jede Bewegung elektrischer Ladungen einen Stromfluß dar, auch die thermische Bewegung der Ladungsträger im Plasma. Beim Fehlen äußerer Einwirkungen bewegen sich aber die Elektronen und Ionen völlig regellos durcheinander, so daß jede Bewegungsrichtung gleichberechtigt ist. Man bezeichnet diese Eigenschaft als *Isotropie* der Geschwindigkeitsverteilung. Sie verhindert, daß durch eine gedachte, beliebig im Plasma liegende Fläche aufgrund der thermischen Bewegung der Ladungsträger ein nachweisbarer Strom fließt. Die von den Elektronen in der

einen Richtung durch die Fläche getragene Ladung entspricht gerade der in Gegenrichtung durch die Fläche hindurchtretenden Ladung. Für die von den Ionen transportierte Ladung gilt natürlich das gleiche. Völlig anders verhält sich das Plasma bei Vorhandensein eines äußeren elektrischen Felds. Der unregelmäßigen Bewegung der Ladungsträger überlagert sich jetzt eine zusätzliche Bewegung in einer bestimmten Richtung, meist in der Richtung des angelegten elektrischen Felds. Dadurch wird die Geschwindigkeitsverteilung der Ladungsträger *anisotrop*. Deshalb treten jetzt durch die gedachte Fläche mehr Ladungsträger in der einen Richtung als in der Gegenrichtung auf. Die von den Elektronen in der einen Richtung durch die Fläche getragenen Ladungen entsprechen nicht mehr den in Gegenrichtung transportierten Ladungen. Es fließt ein von den Plasmaelektronen und ganz entsprechend von den Ionen getragener elektrischer Strom durch das Plasma. Bezeichnen $\boldsymbol{j}$ die Stromdichte, $\boldsymbol{E}$ die treibende elektrische Feldstärke und σ die Leitfähigkeit des Plasmas, so gilt das Ohmsche Gesetz $\boldsymbol{j} = \sigma \boldsymbol{E}$.

Die Elektronen bewegen sich bei diesem Vorgang zusätzlich in Richtung des von außen angelegten elektrischen Felds, die Ionen wegen ihrer entgegengesetzten Ladung in Gegenrichtung. Der Strom setzt sich *additiv* aus den Anteilen beider Ladungsträger zusammen, weil sowohl ihr Ladungsvorzeichen als auch die Richtung ihrer zusätzlichen Bewegung jeweils entgegengesetzt sind. Daher gilt mit $\sigma = \sigma_e + \sigma_i$ die Gleichung $\boldsymbol{j} = \sigma \boldsymbol{E} = (\sigma_e + \sigma_i)\, \boldsymbol{E}$, wenn σ_e der von den Elektronen und σ_i der von den Ionen herrührende Anteil der Leitfähigkeit ist.

Da die zusätzliche gerichtete Geschwindigkeit der Ionen jedoch wegen ihrer geringeren Beweglichkeit (vor allem wegen ihrer vergleichsweise großen Masse) bedeutend kleiner ist als die der Elektronen, kann der von den Ionen übernommene Stromanteil gegenüber dem Anteil der Elektronen meist vernachlässigt werden. Die Gleichung vereinfacht sich daher in vielen Fällen zu $\boldsymbol{j} \approx \sigma_e \boldsymbol{E}$.

Wie in metallischen Leitern wirken die unvermeidlichen Stöße der Ladungsträger untereinander und mit Neutralteilchen dem Stromfluß entgegen. Diese Stöße sind letzten Endes die Ursache dafür, daß die Ladungsträger trotz der anhaltenden Beschleunigung durch das äußere elektrische Feld keine ständige Geschwindigkeitserhöhung erfahren, sondern bald eine konstante Endgeschwindigkeit erreichen. Diese gerichtete Bewegung des Ladungsträgers wird als *Driftbewegung* bezeichnet, die mit ihr verbundene Geschwindigkeit als *Driftgeschwindigkeit* $v_D = \mu E$. Sie ist der wirkenden Feldstärke proportional. Der Proportionalitätsfaktor μ wird *Beweglichkeit* genannt (↑ 8.3.5.3.). Ganz allgemein besteht zwischen der Beweglichkeit μ, der Elementarladung e, der Anzahl n der fraglichen Ladungsträger je Volumeneinheit und der Leitfähigkeit σ der Zusammenhang $\sigma = n\, e\, \mu$. Da sich die Leitfähigkeit σ des Plasmas aus den Anteilen σ_e und σ_i zusammensetzt, gilt $\sigma = e(n_e\mu_e + n_i\mu_i) = \sigma_e + \sigma_i$, wenn n_e, n_i die Anzahlen der Elektronen bzw. der Ionen je Volumeneinheit und μ_e, μ_i die Beweglichkeiten der Elektronen bzw. der Ionen bedeuten. Für die Stromdichte ergibt sich dann $\boldsymbol{j} = e(n_e\mu_e + n_i\mu_i)\, \boldsymbol{E} \approx n_e e \mu_e \boldsymbol{E}$.

Da mit wachsender Temperatur des Plasmas die Stöße seltener werden, ist in heißen Plasmen die elektrische Leitfähigkeit besonders groß; sie kann die der besten metallischen Leiter übersteigen. Diese gute elektrische Leitfähigkeit verhindert schließlich auch, daß in Plasmen über längere Zeiten starke elektrische Felder existieren können. Wegen der großen Beweglichkeit der Ladungsträger, vor allem der Elektronen, ruft jedes stärkere elektrische Feld im Plasma sofort einen Stromfluß hervor, der zur Anhäufung von Überschußladungen an den Grenzflächen des Plasmas führt und dadurch die Stärke des elektrischen Felds im Plasma teilweise reduziert. Dieser Vorgang wird als *Polarisation* bezeichnet (↑ 9.1.3.4.).

9.1.3.4. Ambipolare Diffusion und Randschichten. Ähnlich wie die Atome oder Moleküle im gasförmigen Aggregatzustand versuchen auch die Teilchen des Plasmas eventuell vorhandene lokale Konzentrationsunterschiede durch Diffusion auszugleichen. Da jedoch viele Plasmateilchen elektrische Ladungen tragen, verläuft der entsprechende Diffusionsvorgang wesentlich anders als im gasförmigen Aggregatzustand. Wegen ihrer kleineren Masse und damit größeren Beweglichkeit werden die Elektronen im Konzentrationsgefälle zunächst wesentlich schneller diffundieren als die Ionen. Als Folge davon tritt eine *Ladungstrennung* auf. Die dadurch entstehenden elektrischen Felder wirken der schnelleren Diffusion der Elektronen entgegen, so daß schließlich Elektronen und Ionen *gleich schnell diffundieren*. Diese Erscheinung wird als *ambipolare Diffusion* bezeichnet. Die Verhältnisse liegen ähnlich wie bei der Diffusion von Ladungsträgern in Festkörpern (↑ 8.3.5.1.). Auf ihrer Grundlage vollziehen sich die meisten Diffusionsvorgänge im Plasma.

Die ambipolare Diffusion steht auch mit der Ausbildung von *Plasmarandschichten*, z. B. in der unmittelbaren Nähe von Gefäßwänden, und mit deren spezifischen Eigenschaften in engem Zusammenhang. Da von den (meist nichtleitenden) Gefäßwänden kein elektrischer Strom abfließen kann, müssen die in der Zeiteinheit auf die Wand diffundierenden Elektronen und Ionen zahlenmäßig etwa gleich sein.

Das ist jedoch nur möglich, wenn den viel beweglicheren Elektronen das Erreichen der Wand erschwert wird. Im Zusammenhang mit der ambipolaren Diffusion bildet sich daher zwischen Wand und Plasma ein elektrisches Gegenfeld aus. Dies ist die Ursache dafür, daß das Plasma in einem Gefäß immer ein höheres Potential hat als die Gefäßwand. Die Potentialdifferenz zwischen Wand und Plasma wird als *Kontaktpotentialdifferenz* bezeichnet. Sie beträgt in typischen Laborplasmen mehrere Zehntel Volt bis zu einigen 10 V.

Die Struktur der Randschichten des Plasmas ist außerordentlich kompliziert und bisher noch wenig erforscht. Für den Verlauf des ortsabhängigen Potentials V im Plasma und in der Randschicht haben LANGMUIR und TONKS die sogenannte vollständige Gleichung des Plasmas und der Randschicht angegeben:

$$\text{div grad } V - 4\pi e n_0 \exp\left[qV/(k_B T_q)\right] + 4\pi r^{-\beta} \int_0^r N_z z^\beta \, dz/v_z(V) = 0;$$

in ihr bedeuten n_0 die Elektronendichte im Plasma in sehr großer Entfernung von der Randschicht, e die Elementarladung, k_B die Boltzmannkonstante und T_q die Elektronentemperatur; r und z sind Ortskoordinate und Integrationsvariable, N_z bezeichnet die Anzahl der positiven Ionen, die in der Zeiteinheit je Volumeneinheit an der Stelle z im Plasma gebildet werden, v_z die Geschwindigkeit eines im Punkt z gebildeten Ions an der Stelle r; β schließlich bedeutet eine ganze Zahl, die je nachdem, ob ein ebenes, zylindrisches oder sphärisches Problem vorliegt, zu 0, 1 oder 2 gewählt werden muß. Diese bereits unter sehr vereinfachenden Annahmen aufgestellte Beziehung konnte bisher nur für sehr wenige Spezialfälle gelöst werden.

9.1.3.5. Verhalten des Plasmas im Magnetfeld. Während die geladenen Teilchen des Plasmas längs der Kraftlinien des Magnetfelds frei beweglich bleiben, erfahren sie senkrecht zu den magnetischen Feldlinien eine Kraftwirkung gemäß der Lorentzkraft $\boldsymbol{K} = e(\boldsymbol{v} \times \boldsymbol{B})$, in der e die Elementarladung, $\boldsymbol{v}$ den Vektor der Teilchengeschwindigkeit und $\boldsymbol{B}$ den Vektor der Magnetfeldstärke bezeichnen. Daher können sie sich quer zu den Magnetfeldlinien nur auf kreisförmigen Bahnen bewegen (↑ 2.2.3.). Aus diesen beiden Bewegungsmöglichkeiten resultieren für die geladenen Plasmateilchen schraubenförmige Bahnkurven um die Kraftlinien des Magnetfelds. Besonders kompliziert werden die Verhältnisse noch dadurch, daß die Ladungsträger von Zeit zu Zeit mit anderen Teilchen zusammenstoßen und dabei sprunghafte Richtungsänderungen erfahren. Im ganzen aber wird durch die Wirkung des Magnetfelds die freie Bewegung der Ladungsträger wesentlich eingeschränkt; sie sind gewissermaßen an die magnetischen Kraftlinien *gefesselt* (Abb. 9.1.3-4). Das gilt insbesondere für dünne und heiße Plasmen, in denen der durch das Magnetfeld erzwungene Ordnungszustand der Trägerbewegung nicht durch viele Stöße mit anderen Teilchen gestört wird.

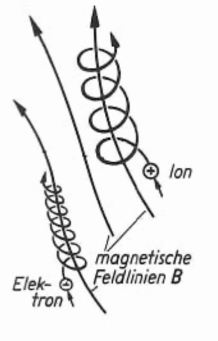


Abb. 9.1.3-4 Schraubenbewegung geladener Teilchen im Magnetfeld

Die Kopplung zwischen Plasma und Magnetfeld ist so innig, daß andererseits auch bei der Bewegung des Plasmas als Ganzes die Magnetfeldlinien verformt und mitgenommen werden. Man spricht geradezu davon, daß Plasma und Magnetfeld sich gegenseitig *einfangen* bzw. *einfrieren*. Größte Bedeutung hat dieses besondere Verhalten des Plasmas im Magnetfeld für die Erzeugung sehr heißer Plasmen. Allein Magnetfelder sind in der Lage, Plasmen bei Temperaturen über 10^5 K längere Zeit für die Untersuchung und weitere Aufheizung festzuhalten. Alle anderen stofflichen Gefäßwände würden bei solchen hohen Temperaturen selbst nur im Plasmazustand existieren können und damit zur *Halterung* eines Plasmas nicht mehr in Frage kommen. Magnetfeldanordnungen, in denen sehr heiße Plasmen erzeugt, aufgeheizt und zusammengehalten werden, bezeichnet man als *Magnetfallen* (↑ 9.4.2.).

9.1.3.6. Schwingungen und Wellen im Plasma. Das Plasma ist ein schwingungsfähiges Gebilde; viele verschiedene Typen von Schwingungen können sich in ihm ausbilden und als Wellen ausbreiten. Abgesehen von den gewöhnlichen *Schallschwingungen* der neutralen Gasatome, die schon im gasförmigen Aggregatzustand möglich sind, bestehen die einfachsten Schwingungsformen aus der geordneten Schwingung der Ionen und Elektronen, die durch spontane oder gewollte Ladungstrennung in kleinsten Bereichen zustande kommt. Je nachdem, ob Ionen oder Elektronen betrachtet werden, spricht man von *elektrostatischen Ionen-* oder *Elektronenschwingungen*. Sie erfolgen ähnlich wie Schallschwingungen *longitudinal* und werden daher oft auch als *Ionen-* oder

Elektronenschall bezeichnet. Die Frequenz der elektrostatischen Elektronenschwingungen, die Plasmaelektroneneigenfrequenz $\omega_e = 2\pi f_e = [n_e e^2/(\pi\varepsilon_0 m_0)]^{1/2}$, hängt allein von der im Plasma vorhandenen Elektronendichte n_e ab; e bedeutet die Elementarladung und m_0 die Elektronenmasse. Für typische Laborplasmen mit $n_e = 10^8$ bis $10^{14}\,\text{cm}^{-3}$ ergeben sich Plasmaelektroneneigenfrequenzen von 10^8 bis 10^{11} Hz; das sind Frequenzen im Höchstfrequenzgebiet.

Solche lokal entstehenden Schwingungen können sich über weite Gebiete des Plasmas als Wellen ausbreiten. Ihre Ausbreitungsgeschwindigkeit hängt im allgemeinen stark von der Frequenz ab. Dies wird als *Dispersionseigenschaft* des Plasmas bezeichnet.

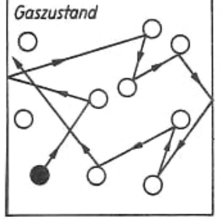
Sind stärkere Magnetfelder vorhanden, so können Schwingungen entstehen, die auf der innigen Kopplung zwischen Magnetfeld und Plasma beruhen. Bei einer Bewegung des Plasmas werden dadurch die Kraftlinien des Magnetfelds mitgenommen (oder eingefroren); dabei entstehen rücktreibende Kräfte, die Schwingungen hervorrufen. Diese können sich ebenfalls längs oder quer zum Magnetfeld als Wellen ausbreiten. Die sich längs des Magnetfelds ausbreitenden Wellen werden als *magnetohydrodynamische* oder *Alfvénwellen* bezeichnet, die sich quer dazu ausbreitenden als *Magnetschallwellen*. Infolge Dispersion und Dämpfung besteht für viele der möglichen Schwingungstypen die Möglichkeit, sich vor allem an Plasmagrenzflächen ineinander umzuwandeln.

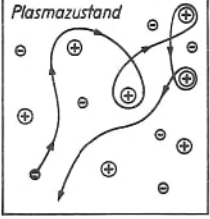
Außer diesen im Plasma selbst erzeugten, für das Plasma spezifischen Schwingungen, kann das Plasma auch außerhalb erzeugte akustische oder elektromagnetische Schwingungen fortleiten. Für *akustische* oder *Schallschwingungen* ist dazu nur eine ausreichende Neutralgasdichte erforderlich; ansonsten verläuft die Ausbreitung solcher Wellen wie im gasförmigen Aggregatzustand. Im Hinblick auf die Ausbreitung *elektromagnetischer Wellen* charakterisiert man das Verhalten des Plasmas am besten durch einen (komplexen) Brechungsindex, dessen hier allein interessierender reeller Anteil $n = \sqrt{1 - f_e^2/f^2}$ beträgt, wenn f_e die Plasmaelektroneneigenfrequenz und f die Frequenz der betrachteten elektromagnetischen Wellen bedeuten. Während im Bereich des festen, des flüssigen und des gasförmigen Aggregatzustands der Brechungsindex immer größer als Null, meist sogar größer als Eins ist und nur wenig von der Frequenz abhängt (↑ 8.8.1.), ist er im Plasmazustand sehr stark frequenzabhängig. Für elektromagnetische Schwingungen, die mit einer der Plasmaelektroneneigenfrequenz entsprechenden Frequenz erfolgen, wird er sogar Null. Daher können sich durch das Plasma nur elektromagnetische Wellen ausbreiten, deren Frequenz größer als die Plasmaelektroneneigenfrequenz ist. Für elektromagnetische Wellen mit kleinerer Frequenz hingegen tritt immer *Totalreflexion* ähnlich wie an einer Metallplatte ein.

Diese Erscheinung spielt für den Kurzwellenfunkverkehr über weite Entfernungen eine große Rolle; sie erklärt die Totalreflexion der Rundfunkwellen an den Ionosphärenschichten.

9.1.3.7. Coulombwechselwirkung, kollektive Erscheinungen, Instabilitäten. Im gasförmigen Aggregatzustand treten die Teilchen nur unmittelbar beim Stoßvorgang selbst in Wechselwirkung miteinander, während sie zwischen den Stößen völlig frei und unbeeinflußt voneinander bleiben. Dieser Umstand führt zu einer geradlinigen Teilchenbewegung, die nur bei Zusammenstößen durch sprunghafte Änderung der Bewegungsrichtung unterbrochen wird. Zwischen den geladenen Teilchen des Plasmas hingegen liegen die Verhältnisse völlig anders. Für die geladenen Teilchen gilt stets das *Coulombsche Gesetz* $F = q_1 q_2/4\pi\varepsilon_0 r^2$, mit F wechselseitig aufeinander ausgeübte Kraft, q_1, q_2 die Ladungen der wechselwirkenden Teilchen und r ihr Abstand. Danach nimmt die Kraftwirkung mit der Entfernung nur schwach ab, so daß praktisch jedes geladene Teilchen ständig mit seinen zahlreichen geladenen Nachbarn in mehr oder minder starker Wechselwirkung steht. Dadurch sind die Bahnen der geladenen Teilchen im Plasma nicht mehr geradlinig, sondern erscheinen ständig mehr oder minder stark gekrümmt (Abb. 9.1.3-5). Bei der kinetischen Behandlung der Bewegung der Ladungsträger im Plasma müssen deshalb im Grunde genommen die Wirkungen aller übrigen geladenen Teilchen berücksichtigt werden. Die damit verbundene Problematik wird als *Vielkörperproblem* der Plasmaphysik bezeichnet und ist exakt nicht lösbar.

Andere dem Plasmazustand eigene sehr kom-







⊖ Elektronen } festgehalten
⊕ Ionen } gedacht
⊖ bewegtes Elektron

Abb. 9.1.3-5 Thermische Bewegung im Gas- und im Plasmazustand

plizierte Wechselwirkungen, die nicht mehr durch Coulombkräfte allein zu erklären und heute noch wenig erforscht sind, werden als *kollektive* oder *kooperative Erscheinungen* bezeichnet. Ihnen wird eine Reihe bisher noch unverstandener seltsamer Eigenschaften des Plasmazustands zugeschrieben. Solche kollektiven Erscheinungen spielen bei der Einschließung und Halterung von heißen Plasmen in *Magnetfallen* (↑ 9.4.2.) eine große Rolle; z. B. ist es bisher noch nicht gelungen, oberhalb bestimmter sehr hoher Temperaturen Plasmen in Magnetfallen stabil zu haltern, da sich dann, wahrscheinlich wegen solcher kollektiver Erscheinungen, gewöhnlich Turbulenzerscheinungen im Plasma ausbilden, die die *magnetischen Wände* der Magnetfallen zerstören und das eingeschlossene Plasma entweichen lassen.
Man bezeichnet Verhaltensweisen des Plasmas, bei denen kleinste zufällige Störungen zu einer allgemein verheerenden Unordnung anwachsen, als *Instabilitäten.*
Zweifellos bildet die noch nicht ausreichend beherrschte Stabilität sehr dichter und heißer Plasmen eines der wichtigsten Probleme auf dem Wege zur gesteuerten Kernfusion mittels Plasmaanlagen (↑ 9.4.2.).

9.1.4. Klassifizierung der Plasmen

Eine Klassifizierung der Plasmen kann nach sehr vielen verschiedenen Gesichtspunkten vorgenommen werden. Physikalisch aufschlußreich jedoch ist allein die Unterscheidung nach ihren inneren Eigenschaften, die durch die inneren *Plasmaparameter* bestimmt werden. Zu diesen gehören vor allem Druck, Temperatur und Ladungsträgerkonzentration (Anzahl der Ladungsträger je cm^3); z. B. unterscheidet man zwischen *Hoch-* und *Niederdruckplasmen*, je nachdem, ob der manometrisch gemessene Druck den Atmosphärendruck über- oder unterschreitet.
Nach dem thermodynamischen Verhalten unterscheidet man zwischen *Gleichgewichts-* und *Nichtgleichgewichtsplasmen*, je nachdem, ob sich zwischen den energieaustauschenden Plasmateilchen ein Gleichgewichtszustand eingestellt hat oder nicht. In einem *isothermischen Plasma* befinden sich die Teilchen aller Teilchensorten miteinander in vollständigem thermodynamischem Gleichgewicht (↑ 9.1.3.1.), in einem nichtisothermischen liegt nur partielles thermodynamisches Gleichgewicht oder gar kein Gleichgewicht vor. Nur für isothermische Plasmen hat der Begriff der *Plasmatemperatur* Bedeutung. Solche Plasmen werden als *kalte Plasmen* bezeichnet, wenn ihre Temperatur unter 10^5 K liegt. Von *heißen Plasmen* spricht man erst bei Temperaturen über 10^6 K.

Eine weitere Unterscheidung richtet sich nach dem *Ionisationsgrad* $x = n_i/(n_i + n_n)$ des Plasmas, dem Verhältnis der in einem Kubikzentimeter enthaltenen Anzahl n_i von Ionen zu der in derselben Volumeneinheit enthaltenen Anzahl $(n_i + n_n)$ von Ionen und Atomen insgesamt. Er wird als echter Bruch oder in Prozenten angegeben. Mit wachsender Temperatur durchläuft der Ionisierungsgrad eines Gases alle Werte von Null bis Eins (Abb. 9.1.4-1). Für

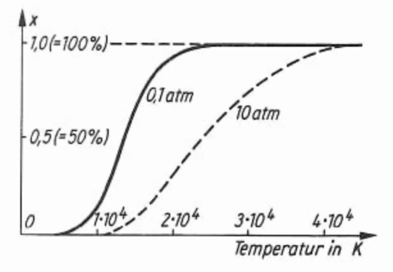


Abb. 9.1.4-1 Temperaturabhängigkeit des Ionisierungsgrads x (qualitativ)

$x = 0$ sind keine Ionen vonrhaden und es liegt ein neutrales oder nichtionisiertes Gas vor; für $x = 1$ sind dagegen alle Atome ionisiert, und man spricht von einem *vollständig ionisierten Plasma.* Normalerweise ist im Plasmazustand nur ein Teil der Atome ionisiert; es liegt dann ein *teilweise* oder *partiell ionisiertes Gas* oder Plasma vor. Je nachdem, ob von den Atomen bei der Ionisation nur ein oder gar mehrere Elektronen abgetrennt sind, unterscheidet man zwischen *einfach* und *mehrfach ionisierten Plasmen.* In einem *Molekülplasma* sind ionisierte Moleküle in beträchtlicher Anzahl vorhanden; ein *Kernplasma* enthält neben den freien Elektronen nur noch nackte Atomkerne.
Nach der Ladungsträgerkonzentration n_e unterscheidet man *dünne Plasmen* mit Elektronenkonzentrationen unter 10^8 cm^{-3} von *dichten Plasmen* mit mehr als 10^{14} Elektronen je Kubikzentimeter. In bezug auf mehr äußerliche Merkmale unterscheidet man z. B. nach dem Zeitverhalten zwischen *stationären Plasmen,* die genügend lange Zeit existieren, und *nichtstationären* oder *Impulsplasmen,* die nur sehr kurzzeitig aufrechterhalten werden können. *Feldfreie Plasmen* enthalten keine oder fast keine elektrischen oder magnetischen Felder, in *magnetfeldbehafteten* oder *Magnetoplasmen* dagegen sind stärkere Magnetfelder vorhanden. Im Grunde genommen ist jedes Plasma, in dem stärker gerichtete elektrische Ströme fließen, magnetfeldbehaftet. Sind die Eigenschaften des Plasmas nicht von der Beobachtungsrichtung abhängig, so heißt es *isotrop,* andernfalls *anisotrop.* Beim Vorhandensein von stärkeren Ma-

gnetfeldern sind die Plasmen in bezug auf viele Eigenschaften gewöhnlich anisotrop. Wenn Neutralgas- und Ladungsträgerkonzentration im Plasma über große Raumbereiche nahezu konstant sind, wird das Plasma als *homogen* bezeichnet, andernfalls als *inhomogen*. Jedes Plasma ist im Bereich seiner Grenz- bzw. Randschichten inhomogen.
Die in diesem Zusammenhang dargelegten Merkmale und Eigenschaften überschneiden sich natürlich teilweise, so daß für ein reales Plasma meist gleichzeitig mehrere der angeführten Bezeichnungen zutreffend sind.

9.1.5. Messungen am Plasmazustand

9.1.5.1. Plasmadiagnostik. Wegen der komplexen Struktur des Plasmas und seines eigenartigen Verhaltens wird jede Messung am Plasmazustand zu einem Problem. Man ist deshalb in der Regel bemüht, die Eigenschaften eines Plasmas auf möglichst wenige, dafür aber grundsätzlich wichtige Kenngrößen zurückzuführen, aus denen die interessierenden Plasmaeigenschaften dann indirekt erschlossen oder errechnet werden können. Zu diesen wichtigen Kenngrößen oder *Plasmaparametern* gehören z. B. der Druck, die Temperatur des Plasmas oder der einzelnen Teilchensorten, die Konzentration der einzelnen Teilchensorten, die elektrische und die magnetische Feldstärke im Plasmainneren. Mit ihrer Bestimmung sowie allen damit zusammenhängenden Problemen beschäftigt sich die *Plasmadiagnostik*, die sich im Laufe der Zeit zu einem speziellen Teilgebiet der Plasmaphysik entwickelt hat.
9.1.5.2. Verschiedene Untersuchungsmöglichkeiten. Der Druck kann am einfachsten, meist ähnlich wie im gasförmigen Aggregatzustand, mit Hilfe von mechanischen Druckmeßeinrichtungen bestimmt werden. Die im Laufe der Zeit entwickelten und vielfach erfolgreich erprobten Meßmethoden für die Bestimmung der meisten anderen Plasmaparameter lassen sich ziemlich zwanglos in einige wesentlich voneinander verschiedene Untersuchungstechniken einordnen. Zu diesen gehören in der Hauptsache die Sondendiagnostik, die optische und spektroskopische Diagnostik und schließlich die Mikrowellendiagnostik. Neben diesen mehr allgemeinen Möglichkeiten gibt es noch sehr viele spezielle Verfahren und Methoden, deren Anwendbarkeit jedoch beschränkt ist und die gewöhnlich nur bei ganz bestimmten Fragestellungen benutzt werden.
9.1.5.2.1. *Sondendiagnostik.* Dieses Verfahren verwendet zum Bestimmen von Plasmaparametern *elektrische Sonden.* Diese bestehen im einfachsten Fall aus winzigen Metallstücken in Platten-, Zylinder- oder Kugelform, die mit einer isolierten Zuleitung versehen und an die zu untersuchende Stelle des Plasmas gebracht werden. Nach ihrem Erfinder werden sie auch als *Langmuirsonden* bezeichnet. Mit ihnen kann das Plasma regelrecht abgetastet werden. Wird ihnen von außen kein elektrisches Potential aufgezwungen, so nehmen sie durch den Aufprall von Elektronen und Ionen aus dem Plasma von selbst ein Potential an, das mit dem Potential des Plasmas am Meßort eindeutig zusammenhängt. Aus dem elektrometrisch gemessenen Potential der Sonde kann daher auf das *Plasmapotential* geschlossen werden. Aus solchen Potentialmessungen an verschiedenen benachbarten Stellen im Plasma lassen sich schließlich Größe und Richtung eines eventuell im Plasmainneren vorhandenen elektrischen Felds ermitteln. Bedeutend wichtigere Informationen über das Plasma können indessen erhalten werden, wenn die Sonde von außen mit einer veränderlichen Spannung beaufschlagt wird. In diesem Fall fließt aus dem Plasma ein elektrischer Strom zur Sonde, der im äußeren Sondenstromkreis gemessen werden kann (Abb. 9.1.5-1). Der Zusammenhang zwischen Sondenstrom und Sondenspannung wird als *Sondencharakteristik* bezeichnet (Abb. 9.1.5-2). Aus dieser können ziemlich weitgehende Rückschlüsse auf bestimmte Plasmakenngrößen gezogen werden.
Diese einfache und informationsergiebige Diagnostikmethode hat im Laufe der Zeit viele Verbesserungen erfahren; z. B. werden

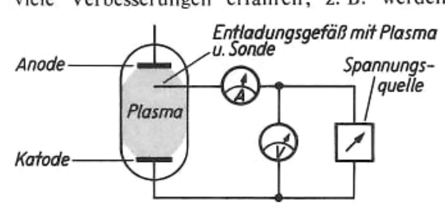


Abb. 9.1.5-1 Sondenmeßkreis

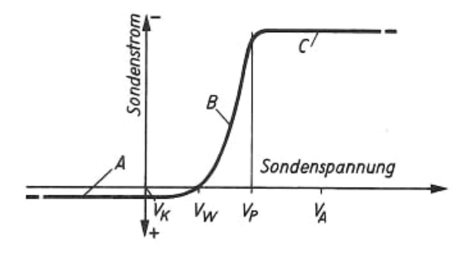


Abb. 9.1.5-2 Sondencharakteristik (qualitativ); *A* Ionensättigungsstrom, *B* Elektronenanlaufstrom, *C* Elektronensättigungsstrom, V_A Anodenpotential, V_K Katodenpotential, V_P Plasmapotential, V_W „Wand"potential

neuerdings vielfach *Gittersonden* verwendet, mit deren Hilfe es sogar möglich war, komplette *Geschwindigkeitsverteilungsfunktionen* (↑ Abb. 9.1.3-1) der im Plasma vorhandenen Elektronen und Ionen aufzunehmen. Schließlich sind in letzter Zeit auch *Hochfrequenz-* oder *Resonanzsonden* in Anwendung gekommen. Diese sind gewöhnlich Langmuirsonden, denen aber zusätzlich eine hochfrequente elektrische Spannung mit veränderlicher Frequenz aufgeprägt wird. Aus der Charakteristik des auf diese Sonde fließenden Hochfrequenzstroms können ebenfalls Rückschlüsse auf Elektronentemperatur, Elektronenkonzentration u. a. gezogen werden. Ein großer Nachteil aller dieser Sondenmethoden besteht allerdings darin, daß die materiellen Sonden in ihrer Nachbarschaft die Eigenschaften des Plasmas mehr oder weniger stark verändern, so daß die aus Sondenmessungen ermittelten Plasmaparameter nicht in jedem Fall die des ungestörten Plasmas sind.

9.1.5.2.2. *Optische und spektroskopische Diagnostik.* Eine der einfachsten optischen Diagnostikmethoden besteht darin, den Durchgang monochromatischen Lichts geeigneter Wellenlänge durch das Plasma zu studieren; z. B. können aus der gemessenen Absorption des Lichts ziemlich direkt die Absorptions- und Emissionseigenschaften des Plasmas ermittelt werden. Des weiteren lassen sich aus der Phasenverschiebung, die die Lichtwellen beim Durchgang durch das Plasma erfahren und die mit Hilfe sogenannter *optischer Interferometer* gemessen wird, der Brechungsindex und die Neutralgaskonzentration des Plasmas bestimmen. Schließlich können aus der Streuung, die Lichtbündel im Plasma erleiden, Elektronenkonzentration und Elektronentemperatur bestimmt werden.

Bedeutend wichtiger sind die spektroskopischen Diagnostikmethoden, bei denen das vom Plasma selbst im infraroten, sichtbaren und ultravioletten Spektralbereich erzeugte Licht mit Hilfe von *optischen Spektrographen* spektral zerlegt wird (↑ 4.3.2.). Da die spektrale Zusammensetzung des in Form von Linien- oder Kontinuumsstrahlung emittierten Lichts (↑ 9.1.2.2.) in starkem Maße von den inneren Eigenschaften des Plasmas abhängt, können aus der gemessenen Intensität und deren Verteilung über das Spektrum tiefgehende Rückschlüsse auf viele Plasmakenngrößen gezogen werden.

Noch weit größere Möglichkeiten ergeben sich für die spektroskopische Diagnostik durch Detailuntersuchungen an ausgewählten einzelnen Spektrallinien, da deren Form und Struktur ebenfalls in starkem Maße von gewissen Plasmakenngrößen abhängen; z. B. erscheinen die Spektrallinien des im Plasma erzeugten Lichts gewöhnlich nicht in ihrer natürlichen Breite und mit der von der Quantentheorie vorbestimmten Wellenlänge. Als Ursache für diese Abweichungen sind der Dopplereffekt, der Starkeffekt und der Zeemaneffekt bekannt. Als *Dopplereffekt* (↑ 2.3.1.) bezeichnet

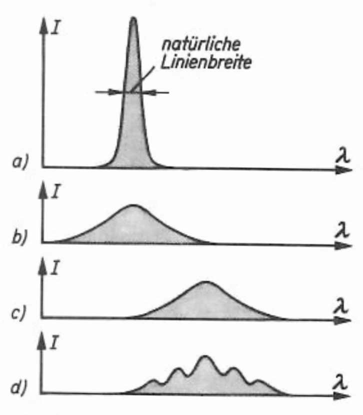


Abb. 9.1.5-3 Verbreiterung, Verschiebung und Aufspaltung von Spektrallinien im Plasma (qualitativ); Intensität I einer Spektrallinie: a) ungestört mit natürlicher Linienbreite, b) durch Stark- und Dopplereffekt verbreitert, c) zusätzlich durch Dopplereffekt verschoben, d) zusätzlich durch Zeemaneffekt aufgespalten; λ Wellenlänge

man die allgemein bekannte Erscheinung, daß jede elektromagnetische oder akustische Strahlung eine Wellenlängen- bzw. Frequenzverschiebung erfährt, wenn sich Strahlungsquelle und Beobachter relativ zueinander bewegen. Da sich die angeregten Atome und Ionen als Quellen der Linienstrahlung im Plasma in ständiger thermischer Bewegung befinden, tritt dieser Effekt auch im Plasma auf. Wegen des statistischen Charakters der thermischen Bewegung führt er summarisch zu einer Verbreiterung der Spektrallinien, der *Dopplerverbreiterung* (Abb. 9.1.5-3). Mit Stark- bzw. Zeemaneffekt bezeichnet man die nur quantentheoretisch zu verstehende Erscheinung, daß viele Spektrallinien in starken elektrischen bzw. magnetischen Feldern in mehrere Komponenten aufspalten (↑ 4.3.3.). Da im Plasma ständig starke elektrische Felder in Form des Mikrofelds (↑ 9.1.3.2.) vorhanden sind, tritt dieser *Starkeffekt* auf und führt wegen des statistischen Charakters des Mikrofelds summarisch zu einer Verbreiterung der Spektrallinien; die *Starkverbreiterung* ist gewöhnlich der Dopplerverbreiterung überlagert. Der *Zeemaneffekt* hingegen kann im Plasma nur auftreten, wenn zusätzlich starke magnetische Felder vorhanden sind. Er führt dann zu einer Aufspaltung

der Spektrallinien, zu der *Zeemanaufspaltung*. Bewegt sich schließlich das Plasma als Ganzes in bezug auf den Beobachter, so ruft der Dopplereffekt noch eine zusätzliche einseitige Verschiebung der ausgesandten Spektrallinien auf der Wellenlängenskala hervor, die als *Dopplerverschiebung* (↑ Abb. 9.1.5-3) bezeichnet wird. Durch die Überlagerung von Dopplerverbreiterung, Starkverbreiterung und u. U. noch der Zeemanaufspaltung ist die genaue Form der Spektrallinien, die man auch als *Spektrallinienprofil* bezeichnet, ziemlich kompliziert. Um die einzelnen genannten Effekte zu plasmadiagnostischen Zwecken ausnutzen zu können, ist es nötig, sie voneinander zu trennen.

Neuerdings setzt man für die Plasmadiagnostik zunehmend die Laser-Fluoreszenz-Spektroskopie ein. Mit Hilfe eines durchstimmbaren Lasers (↑ 4.1.3.) wird Licht einer scharf definierten, jedoch variierbaren Frequenz in das Plasma eingestrahlt. Falls die Quantenenergie mit Anregungsenergien der Teilchen im Plasma übereinstimmt, kommt es zur Elektronenanregung und nachfolgenden Fluoreszenzstrahlung. Infolge der Frequenzschärfe des Laserlichts reichen die Dopplerverschiebungen bereits aus, um eine Verstimmung zu bewirken. Daraus kann auf die Geschwindigkeitsverteilung der Teilchen im Plasma geschlossen werden.

9.1.5.2.3. *Mikrowellendiagnostik.* Als *Mikrowellen* bezeichnet man elektromagnetische Wellen mit Wellenlängen im Zentimeter- und Millimetergebiet. Ähnlich wie bei den optischen und spektroskopischen Methoden wird auch bei den Mikrowellenverfahren entweder der Durchgang solcher Mikrowellen durch das Plasma studiert oder aber der vom Plasma selbst emittierte, im Mikrowellengebiet liegende Anteil der kontinuierlichen Bremsstrahlung zur Plasmadiagnostik ausgenutzt; z. B. können

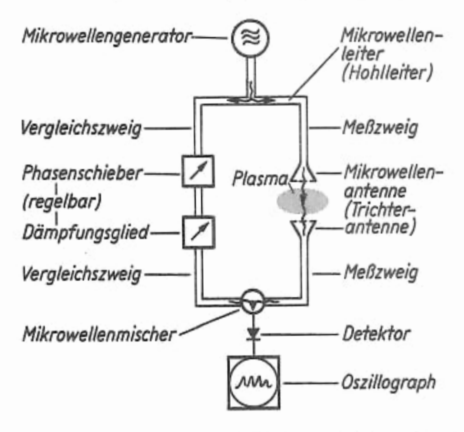


Abb. 9.1.5-4 Schematische Darstellung eines Mikrowelleninterferometers

bei der Durchstrahlung von Plasmen mittels Mikrowellen die Reflexions-, Absorptions- und Transmissionseigenschaften gemessen werden. Besondere Bedeutung kommt jedoch der Phasenverschiebung zu, die die Mikrowellen beim Durchgang durch das Plasma erfahren. Sie kann mittels *Mikrowelleninterferometern* (Abb. 9.1.5-4) bestimmt werden. Aus der Phasenverschiebung lassen sich die Dielektrizitätskonstante und daraus schließlich die Elektronenkonzentration des Plasmas ermitteln. Das vom Plasma selbst erzeugte *Mikrowellenrauschen* kann mittels empfindlicher Empfänger, der *Mikrowellenradiometer*, gemessen und zur Bestimmung der *Rauschtemperatur* verwendet werden, die meist mit der Elektronentemperatur identisch ist, da das Mikrowellenrauschen nichts anderes darstellt als den im Mikrowellengebiet liegenden Anteil der Plasmaelektronenbremsstrahlung (↑ 9.1.2.2.3.).

An genügend dichten Plasmen werden eingestrahlte Mikrowellen reflektiert. Auch aus dieser Erscheinung läßt sich die Elektronenkonzentration bestimmen. Wenn sich die reflektierende Fläche des Plasmas zusätzlich bewegt, erleiden die reflektierten Mikrowellen eine Dopplerverschiebung (↑ 9.1.5.2.2.), aus der die Bewegungsgeschwindigkeit der reflektierenden Plasmafläche bestimmt werden kann. Diese Methode zur Lokalisierung und Verfolgung sich bewegender Plasmaschichten wird als *Dopplerradar* bezeichnet. Schließlich erfahren Mikrowellen in magnetfeldbehafteten Plasmen längs der Kraftlinien eine Drehung ihrer Polarisationsebene, eine Erscheinung, die aus der Optik als *Faradayeffekt* bekannt ist. Die Größe dieser Drehung hängt von der Elektronenkonzentration und der Magnetfeldstärke ab und kann deshalb zur Bestimmung dieser Parameter benutzt werden.

9.1.5.2.4. *Weitere spezielle Diagnostikmethoden.* Eine oft genutzte Möglichkeit zum Studium von schnellen Bewegungsvorgängen an und in Plasmen stellt die *Hochgeschwindigkeitskinematographie* dar. Bei dieser werden in sehr schneller zeitlicher Folge fotografische Aufnahmen des Plasmas vorgenommen. Moderne hochgeschwindigkeitskinematographische Einrichtungen ermöglichen mehr als 10^6 einzelne Aufnahmen je Sekunde. Bei einfacheren Einrichtungen verzichtet man von vornherein auf die Aufnahme von Einzelbildern und zieht statt dessen den gesamten zeitlichen Vorgang räumlich in Form eines Schmierbilds auseinander.

Bei *Elektronenstrahlsonden* macht man die Ablenkung und die Aufweitung, die durch das Plasma geschossene Strahlen schneller Elektronen erfahren, auf einem hinter dem Plasma angeordneten Leuchtschirm sichtbar und schließt

daraus auf die Größe elektrischer oder magnetischer Felder im Plasma. Ebenso können *γ-Strahlen* oder *Korpuskularstrahlen*, z. B. α-Strahlen, durch das Plasma geschossen werden. Aus ihrer Intensitätsabnahme lassen sich dann ebenfalls gewisse innere Eigenschaften ermitteln, z. B. die Neutralgasdichte. Schließlich ist es möglich, durch Zugabe *radioaktiver Beimengungen* die im Plasma ablaufenden Diffusionsvorgänge zu studieren. In ähnlicher Weise können durch Zugabe von *Kohlenstoffteilchen*, die im heißen Plasma aufglühen, makroskopische Bewegungsvorgänge im Plasma sichtbar gemacht werden.

Zur Ausmessung eventuell im Plasma vorhandener veränderlicher Magnetfelder werden vielfach *Magnetfeldsonden* verwendet, die aus winzigen Drahtspulen bestehen, in denen das durchgreifende zeitveränderliche Magnetfeld eine elektrische Spannung induziert. In ähnlicher Weise läßt sich in stromführenden Plasmen die elektrische Stromdichte mit *Rogowskispulen* bestimmen. Das sind winzige toroidförmige Drahtspulen, in denen der durch die Toroidfläche hindurchtretende Strom über sein Magnetfeld eine elektrische Spannung induziert, aus der sich dann die Stromdichte am Meßort bestimmen läßt.

Neben diesen Untersuchungsmethoden werden in der Plasmadiagnostik auch noch Verfahren und Techniken aus anderen Arbeitsgebieten verwendet; z. B. gewinnt die *Massenspektrometrie* in der Plasmadiagnostik immer mehr an Bedeutung. Mit ihrer Hilfe können so wichtige Fragen wie Zusammensetzung des Plasmas nach Gasart, Ladungszustand der Ionen u. a. sehr direkt und eingehend studiert werden. Schließlich spielen bei der Diagnostik sehr heißer Plasmen, in denen Kernreaktionen auftreten können, auch *kernphysikalische Meßmethoden* eine große Rolle.

9.1.6. Vorkommen und Erzeugung von Plasmen

9.1.6.1. Vorkommen im Kosmos und in der höheren Erdatmosphäre. Während Plasmen unter irdischen Bedingungen recht selten anzutreffen sind und ziemlich ungewöhnlich erscheinen, befindet sich im Kosmos der überwiegende Teil der Materie im Plasmazustand. Hier ist das Auftreten der Stoffe in kondensierten Phasen, im festen, flüssigen oder gasförmigen Aggregatzustand, geradezu eine Ausnahme. Sowohl in den *Fixsternen* als auch in der *interstellaren Materie* tritt der Plasmazustand auf (↑ 10.3.3. und 10.3.4.). Besonders der physikalische Zustand der Sonne ist ausführlich untersucht worden (↑ 10.2.2.). Auch die Strahlungsgürtel der Erde, die Polarlichter und die Ionosphäre lassen sich ohne die Gesetze der Plasmaphysik nicht erklären (↑ 10.1.3.).

9.1.6.2. Vorkommen auf der Erde. In geringerem Maße als im Kosmos und in den höheren Atmosphärenschichten sind natürliche Vorkommen des Plasmazustands an der Erdoberfläche oder in ihrer unmittelbaren Nähe zu finden. Eines der eindruckvollsten Beispiele bilden die *Blitzerscheinungen* bei Gewittern, bei denen in der sehr kurzen Zeit von 10^{-5} bis 10^{-6} s infolge eines *elektrischen Durchschlags* die Luft im Blitzkanal ionisiert wird (Tafeln 50/51). Der entstehende Plasmaschlauch kann sich bei einem Durchmesser von nur wenigen Zentimetern bis zu Dezimetern über eine Länge von mehreren Kilometern erstrecken. Obwohl die Durchbruchsspannungen im allgemeinen mehrere hundert Kilovolt betragen und im Plasmakanal Ströme von 10^4 bis 10^5 Ampere fließen, sind die in Blitzentladungen enthaltenen elektrischen Energien wegen der außerordentlichen Kürze der einzelnen Entladungsstöße relativ klein, gewöhnlich nur Bruchteile einer Kilowattstunde.

Eine andere allgemein bekannte Erscheinungsform des Plasmazustands tritt uns in den *Flammenerscheinungen* fast aller Verbrennungsvorgänge entgegen. Ähnliches gilt für die sehr heißen und damit leuchtenden *Auspuffgase* von Kraftmaschinen, von Verbrennungsmotoren, Strahltriebwerken u. a. Dabei dokumentiert sich der Plasmazustand der Flammen- bzw. Auspuffgase hauptsächlich in der großen elektrischen Leitfähigkeit.

Auch die bei *elektrischen Kurzschlüssen* auftretenden Lichterscheinungen sind mit dem Plasmazustand verbunden. Es handelt sich hierbei gewöhnlich um wenig definierte, kurzzeitige Bogenentladungen (↑ 9.1.7.3.3.), die zwischen den spannungsführenden Stromleitern kurz vor der Berührung auftreten und nach Aufhebung der Berührung noch kurzzeitig andauern. Dabei befinden sich die in der Nähe der Kurzschlußstelle vorhandene Luft und abgedampftes Leitermaterial im Plasmazustand.

Eine weitere in der Natur zu beobachtende Plasmaerscheinung besteht in den gelegentlich an Hochspannungsleitungen auftretenden *Koronaentladungen*. Sie entstehen, wenn unter ungünstigen Bedingungen die elektrische Feldstärke zwischen den Hochspannungsleitern so hoch ist, daß zufällig vorhandene Elektronen die umgebende Luft durch Stoß ionisieren. Ähnliche Plasmaerscheinungen stellen die in der Natur mitunter zu beobachtenden *Spitzen-* und *Büschelentladungen* dar, die im Volksmund unter der Bezeichnung Elmsfeuer bekannt sind.

Nach den heutigen Vorstellungen der Wissenschaft muß sich auch in größeren Tiefen der Erde die Materie im Plasmazustand befinden. Die Ionisation dürfte sich dort wegen des im Erdinneren herrschenden enormen Drucks auf

der Grundlage der *Druckionisation* vollziehen. Wegen der im Erdkern vorliegenden hohen Dichten ähnelt das *Plasma des Erdinneren* jedoch mehr einem Metall als einem ionisierten Gas. Es gehört daher nicht unmittelbar in den Bereich der Plasmaphysik.

9.1.6.3. Plasmaerzeugung durch Erwärmen, thermische Ionisierung. Die wichtigsten Methoden der künstlichen Erzeugung von Plasmen beruhen auf der Ionisierung durch Elektronenstoß (↑ 9.1.1.3.1.). Dieses Verfahren wird sowohl bei der Plasmaerzeugung mittels Erwärmung, der *thermischen Ionisierung*, als auch bei der Plasmaerzeugung in *elektrischen Gasentladungen* ausgenutzt.

Die zur Ionisierung erforderliche Energie wird dem Gas von außen durch Erwärmung zugeführt. Die bei Stößen der Gasteilchen mit den künstlich erhitzten Gefäßwänden übertragene Energie ruft eine immer heftigere thermische Bewegung *aller* Teilchen hervor. Die dabei von zufällig vorhandenen *Elektronen* erworbenen Energien reichen schließlich aus, Atome und Moleküle zu ionisieren. Es entstehen neue Elektronen, die schnell ausreichende kinetische Energien gewinnen und ihrerseits weitere Ionisationsakte durchführen können. Bei sehr hohen Temperaturen schließlich führen auch Stöße der Atome und Ionen untereinander direkt zur Ionisierung. Da bei dieser Art der Plasmaerzeugung alle Gas- bzw. Plasmateilchen ihre Energie letzten Endes aus Stößen mit den heißen Gefäßwänden schöpfen, befinden sich alle Teilchen im vollständigen thermodynamischen Gleichgewicht miteinander (↑ 9.1.3.1.). Das entstehende Plasma ist daher streng isotherm, man kann ihm eine einheitliche Plasmatemperatur zuordnen und thermodynamische Vorstellungen anwenden.

Durch den Vorgang der Rekombination ist der erreichbaren Ionisierung eine Grenze gesetzt. Der sich im stationären Plasma (↑ 9.1.4.) zwischen Ionisation und Rekombination ausbildende Gleichgewichtszustand hängt im allgemeinen in komplizierter Weise von den im Plasma ablaufenden Elementarprozessen ab. Im hier interessierenden Fall der thermischen Ionisierung führt die der fortlaufenden Ionisierung entgegenwirkende Rekombination schließlich zu einem Gleichgewichtszustand, der durch die sogenannte *Sahaformel* beschrieben wird

$$\frac{x^2 p}{1 - x^2} = \left(\frac{2\pi m_0}{h^2}\right)^{3/2} (kT)^{5/2}\, \mathrm{e}^{-E_i/k_B T}.$$

(E_i Ionisierungsenergie, h Plancksches Wirkungsquantum, k_B Boltzmannkonstante, m_0 Elektronenmasse, p Druck im Plasma, T Plasmatemperatur). Dabei gilt $0 < x < 1$.

Diese Beziehung entspricht dem aus der physikalischen Chemie bekannten Massenwirkungsgesetz. Für kleine Ionisationsgrade $x \ll 1$, wie sie bei thermischer Ionisierung gewöhnlich zu erwarten sind, kann die Sahagleichung vereinfacht werden zu

$$x \approx p^{-1/2} \left(\frac{2\pi m_0}{h^2}\right)^{3/4} (k_B T)^{5/4}\, \mathrm{e}^{-E_i/2k_B T};$$

sie erlaubt dann, den interessierenden Ionisierungsgrad x auf relativ einfache Weise aus der Temperatur T und dem Druck p des Plasmas zu bestimmen.

Da selbst die am leichtesten zu ionisierenden Stoffe erst bei einigen 1000 K merklich in den Plasmazustand übergehen, müssen bei dieser Plasmaerzeugungsmethode an den Werkstoff des zu erhitzenden Gefäßes höchste Anforderungen gestellt werden. Am weitesten kommt man mit dem *Kingschen Ofen*, der aus einem hochtemperaturfesten, elektrisch beheizten Kohlerohr besteht (Abb. 9.1.6-1). In ihm lassen sich ohne Zerstörung Temperaturen bis nahezu 4000 K erzielen. Diese Temperaturen reichen jedoch nur aus, um einige leicht ionisierbare Gase, z. B. Caesium, Kalium oder Natrium, merklich in den Plasmazustand überzuführen. Für die meisten anderen Stoffe ist eine thermische Ionisierung im Laboratorium nicht möglich.

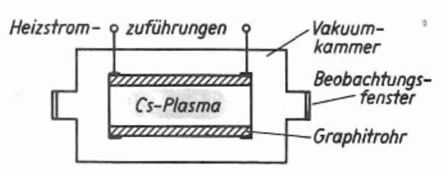


Abb. 9.1.6-1 Prinzipieller Aufbau des Kingschen Ofens

9.1.6.4. Plasmaerzeugung durch elektrische Gasentladungen. Anstatt die Bewegungsenergie aller Gas- bzw. Plasmateilchen durch Temperaturerhöhung im Ofen zu steigern, kann die zur Ionisierung nötige Energie von vornherein vorzugsweise auf die allein sehr ionisierungswirksamen Elektronen übertragen werden. Das geschieht in den elektrischen Gasentladungen mit Hilfe elektrischer Felder. Die im Gas anfangs zufällig vorhandenen Elektronen werden im elektrischen Feld so stark beschleunigt, daß sie ionisieren können. Die dabei entstehenden neuen Elektronen erwerben im elektrischen Feld ebenfalls Energie und können bald weitere Ionisierungen veranlassen. Ein so erzeugtes Plasma, durch das infolge des elektrischen Felds zusätzlich ein elektrischer Strom fließt, bezeichnet man als *Gasentladungsplasma*. Mit dieser Methode sind ohne große Schwierigkeiten praktisch alle Stoffe in den Plasmazustand überführbar. Wegen der leichten und sauberen Erzeugbarkeit der Gasentladungsplasmen und

der Möglichkeit der ziemlich willkürlichen Wahl ihrer Parameter spielen sie für Forschung und Technik eine vorherrschende Rolle.

9.1.6.5. Weitere Möglichkeiten der Plasmaerzeugung. Mehr oder weniger gewollt kann der Plasmazustand auch bei einer Reihe weiterer Vorgänge auftreten; z. B. befindet sich bei allen mit Flammenerscheinungen verbundenen *chemischen Explosions- und Verbrennungsvorgängen* praktisch immer ein Teil der Materie im Plasmazustand. Auch an den Fronten starker Stoßwellen, die sich bei heftigen Detonationen in verdünnten Gasen ausbilden können, ist mitunter der Plasmazustand zu beobachten.

Die Entstehung und Aufrechterhaltung von Plasmen durch Fotoionisierung findet in größtem Maßstab in kosmischen Bereichen statt. Im Laboratorium hat sie bisher nur wenig Verwendung gefunden.

Durch einen starken Stromstoß kann ein dünner Metalldraht in Millisekunden verdampfen. Diese Methode der Drahtexplosion läßt sich benutzen, Plasmen verschiedenster Stoffe dadurch zu gewinnen, daß ein dünner Flüssigkeitsfaden einer Lösung auf eine Elektrode gespritzt wird und durch einen starken Kurzschlußstrom verdampft. Man erhält Plasmen bis zu Drücken von 1 MPa.

Eine recht ungewöhnliche, jedoch schon häufig erprobte Art der Plasmaerzeugung stellt die *Einschubmethode* dar, bei der die durch Beschleuniger auf sehr hohe Energien beschleunigten Ionen in eine Magnetfalle (↑ 9.4.2.) eingeschossen werden. Da die positiv geladenen Ionen aus der Umgebung zufällig vorhandene freie Elektronen anziehen, entsteht auf recht einfache Weise ein sehr heißes Plasma.

9.1.7. Elektrische Gasentladungen

Unter einer elektrischen Gasentladung versteht man einen Vorgang, bei dem ein Gas durch äußere elektrische Einwirkungen in den Plasmazustand übergeführt wird. Die Anordnungen, in denen elektrische Gasentladungen gewöhnlich erzeugt werden, heißen *Gasentladungsröhren*. Sie bestehen im einfachsten Fall aus

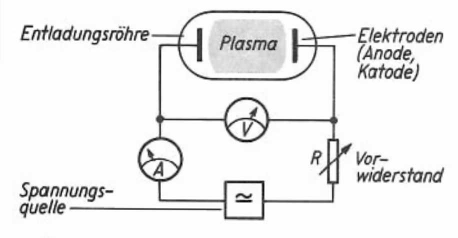


Abb. 9.1.7-1 Gasentladungsstrecke mit äußerem Stromkreis

einem kugel- oder zylinderförmigen mit dem zu ionisierenden Gas gefüllten Glasgefäß, in das zwei Metallelektroden eingeführt sind, die durch eine äußere Stromquelle gespeist werden. Der Raum zwischen den Elektroden, in dem sich die elektrische Gasentladung abspielt, wird als *Gasentladungsstrecke* bezeichnet (Abb. 9.1.7-1).

9.1.7.1. Entladungsmechanismus. Bei sehr kleinen Spannungen zwischen den Elektroden folgen die in sehr geringer Anzahl zufällig im Neutralgas vorhandenen Elektronen dem Zug der elektrischen Kraftlinien und bewirken einen fast unmeßbar kleinen Stromfluß durch das Gas. Durch ständige Zusammenstöße mit den Gasatomen geben diese Elektronen ihre im elektrischen Feld erworbene Energie zum größten Teil wieder ab, so daß die ihnen verbleibende Energie nicht ausreicht, Gasatome zu ionisieren und das Gas in den Plasmazustand überzuführen. Steigert man die Spannung zwischen den Elektroden auf größere Werte, so übersteigt schließlich die von den Elektronen zwischen zwei Zusammenstößen mit Neutralteilchen aus dem elektrischen Feld aufgenommene Energie die Ionisierungsenergie; die Elektronen können durch Zusammenstoß Gasatome ionisieren. Die dabei neu entstandenen Elektronen nehmen aber ebenfalls aus dem elektrischen Feld Energie auf und können ihrerseits weitere Ionisationsakte verursachen. Es bildet sich auf diese Weise sehr schnell eine *Elektronenlawine* aus (↑ Abb. 9.1.1-3), und die Anzahl der Ladungsträger steigt sprunghaft an. Wenn die Ladungsträgerbilanz positiv wird, d. h., wenn je Zeiteinheit mehr Ladungsträgerpaare (Elektronen und Ionen) gebildet werden als durch Rekombinationsvorgänge wieder verschwinden, zündet die Entladung. Die Spannung, bei der dies gerade der Fall ist, wird *Zündspannung* genannt.

Durch die zahlreichen jetzt in der Entladungsstrecke vorhandenen Ladungsträger steigt der Strom durch das Gas sprunghaft auf einen großen Wert an, auf die eigentliche *Entladungsstromstärke*. Gleichzeitig treten eine Reihe von *Sekundärprozessen* in Erscheinung; z. B. schlagen die auf die negative Elektrode, die Katode, auffallenden positiven Ionen dort neue Elektronen heraus. Vielfach wird die Katode durch die auffallenden Ionen auch stark aufgeheizt, so daß aus ihr weitere Elektronen thermisch emittiert werden können Schließlich kann auch die auf die Katode auffallende Ultraviolettstrahlung der Gasentladung Elektronen auslösen. Alle diese zusätzlich an der Katode erzeugten Elektronen, die *Sekundärelektronen*, nehmen nun weiterhin am Entladungsvorgang teil und üben damit einen starken Einfluß auf den Entladungsmechanismus aus. Durch ihr Auftreten und den lawinenhaften Charakter der Elektronenstoßionisierung ist die hohe Zündspannung,

die zur Ingangsetzung des Entladungsvorgangs nötig war, für die weitere Aufrechterhaltung der Entladung nicht mehr erforderlich. Daher kann die Entladung dann mit einer bedeutend niedrigeren Spannung, der *Brennspannung*, betrieben werden.

Einem weiteren Anstieg der Entladungsstromstärke und einem weiteren Abfall der Brennspannung wirken schließlich die unvermeidlichen Verluste an Ladungsträgern infolge Rekombination und Ausscheiden an Wänden und Elektroden entgegen. Gewöhnlich stellt sich sehr bald ein *stationärer Entladungszustand* ein. Die notwendige Bedingung dafür besteht darin, daß *jedes* in der Entladung einmal vorhandene Elektron während seiner Aufenthaltsdauer in der Entladungsstrecke durch Ionisations- und Anregungsakte so viele Ionen und Lichtquanten erzeugen muß, wie zur Auslösung gerade *eines* neuen Sekundärelektrons aus der Katode erforderlich sind. Jedes Elektron muß gewissermaßen für seinen eigenen Ersatz sorgen. Es hat sich auf diese Weise eine *selbständige Entladung* ausgebildet. Im Unterschied dazu bedürfen *unselbständige Entladungen* zu ihrer Aufrechterhaltung weiterer äußerer Mittel, z. B. der zusätzlichen Zufuhr von Elektronen oder Lichtquanten. Bei genügend hohen Spannungen geht jede unselbständige Entladung in eine selbständige über, da dann das Problem der Elektronennachlieferung von der Entladung durch die Sekundärprozesse an der Katode von selbst gelöst wird.

9.1.7.2. Eigenschaften elektrischer Gasentladungen. Der Zusammenhang zwischen Strom und Spannung wird durch die *Strom-Spannungs-Charakteristik* wiedergegeben (Abb. 9.1.7-2). Während in metallischen Leitern die Charakteristik ansteigt und der Zusammenhang zwischen Strom und Spannung dem Ohmschen Gesetz gehorcht, ist die Charakteristik von Gasentladungen über weite Strom- und Spannungsbereiche *fallend*. Diese Eigenschaft ist den elektrischen Gasentladungen ganz allgemein eigen und letzten Endes auf die Ladungsträgervermehrung durch Ionisation zurückzuführen. Durch die über weite Bereiche fallende Charakteristik besteht bei elektrischen Gasentladungen die Möglichkeit, daß sich die Entladungsstromstärke unter Absinken der Brennspannung auf einen das Entladungsgefäß zerstörenden Wert erhöht. Dieser Gefahr begegnet man durch Einschalten eines strombegrenzenden *Vorwiderstands* in die Zuleitungen des Entladungskreises (↑ Abb. 9.1.7-1).

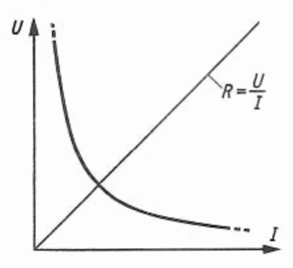


Abb. 9.1.7-2 Fallende Charakteristik einer elektrischen Gasentladungsstrecke; Spannung U, Stromstärke I, Widerstand R

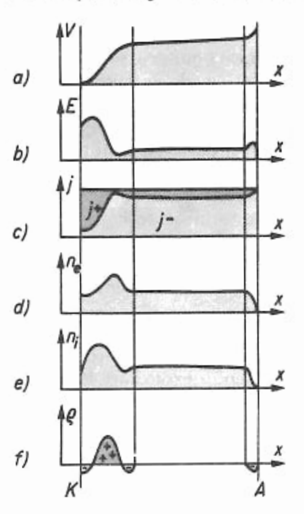


Abb. 9.1.7-3 Glimmentladung: Abhängigkeit a) des Potentials V, b) der elektrischen Feldstärke E, c) der Stromdichte $j = j_+ + j_-$, d) der Elektronenkonzentration n_e, e) der Ionenkonzentration n_i und f) der Raumladung $\varrho = e(n_i - n_e)$ von der Ortskoordinate x zwischen Katode K und Anode A

Die eine bestimmte elektrische Gasentladung charakterisierenden Größen wie Zündspannung, Brennspannung und Entladungsstromstärke werden als *äußere Parameter* bezeichnet. Diese hängen stark von Gasart, Gasdruck, Elektrodenmaterial und Entladungsstreckengeometrie ab. Dadurch ist die Vielfalt der möglichen Gasentladungen außerordentlich groß. Trotzdem weisen die elektrischen Gasentladungen in bezug auf ihre innere Struktur gewisse gemeinsame Züge auf. Besonders deutlich zeigt sich dies im örtlichen Verlauf von *Potential, elektrischer Feldstärke, Raumladung* und *Stromverteilung* (Abb. 9.1.7-3); z. B. befindet sich in unmittelbarer Nähe der Katode praktisch immer ein Gebiet mit starkem Potentialanstieg bzw. mit hoher elektrischer Feldstärke. Dieses wird als *Katodenfallraum*, der in ihm enthaltene Potentialanstieg als *Katodenfall* bezeichnet. Daran schließt sich ein meist ausgedehntes Gebiet an, in dem nur ein

schwacher Potentialanstieg stattfindet bzw. eine nur geringe und praktisch konstante Feldstärke vorhanden ist. Dieses Gebot wird *Rumpf* oder *positive Säule* der Entladung genannt. Schließlich findet sich in unmittelbarer Nähe der Anode, der positiven Elektrode, mitunter ein weiterer kleiner Potentialanstieg oder -abfall, der als *Anodenfall* bezeichnet wird und eine etwas veränderte elektrische Feldstärke hat.

Einen ähnlichen örtlichen Verlauf weisen die in elektrischen Gasentladungen stellenweise stets vorhandenen *Raumladungen* auf. Allein das Gebiet der positiven Säule bleibt weitgehend raumladungsfrei und birgt daher ein echtes quasineutrales Plasma. Schließlich wird der elektrische Strom durch die Entladungsstrecke gebietweise bevorzugt von der einen oder anderen Ladungsträgersorte (Elektronen oder Ionen) getragen. Während in der raumladungsfreien positiven Säule hauptsächlich die Elektronen den Stromtransport übernehmen, sind im Katodenfallgebiet mit seiner starken positiven Raumladung auch die Ionen erheblich daran beteiligt. Die Summe der auf die Elektronen und Ionen entfallenden Stromanteile muß natürlich in jedem Querschnitt der Entladung konstant und der Entladungsstromstärke gleich sein.

Je nach Gasart, Gasdruck, Elektrodenmaterial, Entladungsstreckengeometrie und Entladungsstromstärke sind die einzelnen Entladungsgebiete unterschiedlich gut ausgebildet und verschieden ausgedehnt. Während das Gebiet der positiven Säule vielfach den Hauptteil der Entladungsstrecke einnimmt, sind Katoden- und Anodenfallgebiet gewöhnlich auf die unmittelbare Nachbarschaft der Elektroden beschränkt (Abb. 9.1.7-4). Ungeachtet dessen sind allein die *katodennahen Entladungsteile* für die elektrische Gasentladung notwendig.

9.1.7.2.1. *Katodenfallraum.* Dieses Gebiet schließt sich unmittelbar an die Katode an. Durch den in ihm liegenden starken Potentialanstieg werden einmal die an der Katode erzeugten Sekundärelektronen auf eine zur Ionisierung ausreichende Energie beschleunigt; zum anderen erhalten in ihm die in Gegenrichtung laufenden positiven Ionen eine so große Geschwindigkeit, daß sie beim Auftreffen auf die Katode ausreichend neue Elektronen auslösen. Daher ist die Existenz dieses Katodenfallgebiets mit seinem hohen Potentialanstieg für die Aufrechterhaltung der Entladung notwendig. Da die Elektronen auf ihrem Weg durch den Katodenfallraum selbst nur wenige Ionisierungs- und Anregungsakte durchführen können, erscheint dieses Gebiet relativ dunkel und hebt sich deutlich von den übrigen Teilen der Entladung ab. Wegen der im Katodenfallgebiet vorhandenen starken Raumladungen kann das in ihm enthaltene ionisierte Gas nicht als ein echtes quasineutrales Plasma angesprochen werden. Zum Katodenfallgebiet wird gewöhnlich noch ein hell leuchtendes Gebiet, das *negative Glimmlicht*, hinzugerechnet, das sich durch einen weiteren Dunkelraum von der anschließenden positiven Säule deutlich abhebt. Auch dieses ist für manche Gasentladungen notwendig.

9.1.7.2.2. *Positive Säule.* Sie schließt sich anodenseitig an die katodennahen Entladungsteile an und erstreckt sich meist über weite Gebiete der Entladungsstrecke. Das Gebiet enthält praktisch keine Raumladungen; das in der Säule enthaltene ionisierte Gas bildet so ein echtes quasineutrales Plasma. Unter den Bedingungen der positiven Säule reicht nur bei relativ wenig Elektronen die kinetische Energie zur Ionisierung aus. Um so mehr Elektronen können aber Anregungsprozesse verursachen.

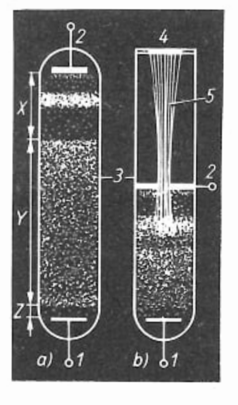


Abb. 9.1.7-4 a) Erscheinungsbild der Glimmentladung; *1* Anode, *2* Katode, *3* gasgefülltes Entladungsrohr, *X* Katodenfallraum (bestehend aus Katodendunkelraum, helleuchtendem negativem Glimmlicht und dem Faradayschen Dunkelraum), *Y* positive Säule (stärker leuchtend), *Z* Anodenfallraum; b) schematische Darstellung einer Kanalstrahlentladung; *4* Leuchtschirm (fluoresziert beim Auftreffen der Kanalstrahlen), **5** Kanalstrahlen

Deshalb ist für die positive Säule eine starke Lichtemission charakteristisch. Für den Entladungsmechanismus selbst ist die positive Säule ziemlich bedeutungslos, sie dient der Entladung hauptsächlich nur zum kontinuierlichen Stromtransport zur Anode; z. B. kann die positive Säule durch Annäherung der Elektroden in ihrer Länge verändert oder gar ganz zum Verschwinden gebracht werden, ohne daß die anderen Entladungsgebiete eine tiefergehende Beeinflussung erfahren.

9.1.7.2.3. *Anodenfallgebiet.* Es bildet sich zwischen positiver Säule und Anode aus und ist, ähnlich wie die positive Säule, für das Zustandekommen und für die Aufrechterhaltung der Entladung ziemlich unwesentlich. In vielen Fällen kann das Anodenfallgebiet als eine ganz gewöhnliche Randschicht aufgefaßt werden, wie sie sich immer an der Grenzfläche eines Plasmas, der positiven Säule, mit einer metallenen Wand, der Anode, ausbildet. Da auch im Anodenfallgebiet Raumladungen vorhanden sind, kann man das darin enthaltene ionisierte Gas nicht als echtes quasineutrales Plasma ansprechen.

9.1.7.3. Einteilung elektrischer Gasentladungen. Eine wesentliche Rolle für das Erscheinungsbild der Gasentladungen spielen das Auftreten von Raum- und Wandladungen, das Aufheizen von Plasma und Elektroden sowie die Wirkung innerer Magnetfelder. Da alle diese Erscheinungen auf den Entladungsstrom zurückgeführt werden können, ist eine Einteilung der elektrischen Gasentladungen nach der Entladungsstromdichte zweckmäßig. Man unterscheidet grob zwischen Dunkel-, Glimm-, Bogen- und Höchststromentladungen, zwischen denen es Übergänge gibt und zu denen noch einige Sonderformen hinzukommen.

9.1.7.3.1. *Dunkelentladungen* werden die Gasentladungen genannt, die nur sehr wenig Licht ausstrahlen; ihre Stromdichte ist so gering, daß sich der Plasmazustand nur wenig ausbilden kann. Bei kleineren Spannungen können sie nur mit äußeren ionisierenden Mitteln aufrechterhalten werden. Sie sind dann unselbständige Entladungen, und ihre Strom-Spannungs-Charakteristik weist keine fallenden Teile auf. Erst bei größeren Spannungen werden sie zu selbständigen Entladungen; ihre Charakteristik ist dann ebenfalls fallend.

9.1.7.3.2. *Glimmentladungen.* Dazu werden alle Gasentladungen gezählt, die bei kleineren Drücken mit relativ geringer Stromdichte ohne äußere ionisierende Mittel brennen; sie sind demnach selbständige Entladungen und ihre Strom-Spannungs-Charakteristik ist über größere Strombereiche fallend. Die zur Aufrechterhaltung der Entladung notwendigen Sekundärelektronen werden in der Regel an der Katode durch schnelle Ionen, die aus dem Entladungsraum auffallen, und in geringerem Maße durch ultraviolette Lichtquanten ausgelöst. Gewisse Sonderformen der Glimmentladungen werden unter den Bezeichnungen *subnormale* und *anomale Glimmentladungen* geführt. Auch die bekannten *Kanalstrahlentladungen* sind eine Sonderform der Glimmentladung. Bei ihnen können die im Katodenfallraum stärker beschleunigten Ionen als Kanalstrahlen, als Strahlen positiver Ionen, durch Löcher in der Katode aus dem Entladungsraum austreten (↑ Abb. 9.1.7-4 b).

9.1.7.3.3. *Bogenentladungen.* Gewöhnlich brennen diese Gasentladungen bei höheren Drücken mit relativ großen Stromdichten. Dadurch wird die Katode stark erhitzt und kann dann Elektronen thermisch emittieren. Deshalb sind zur Aufrechterhaltung solcher Entladungen keine äußeren ionisierenden Mittel erforderlich. Es handelt sich daher um selbständige Entladungen mit fallender Charakteristik. Wegen des hohen Drucks ist der Katodenfallraum auf eine sehr dünne Schicht in unmittelbarer Nähe der Katode beschränkt und gewöhnlich nur schwer feststellbar. Die in ihm beschleunigten Elektronen können kaum durch direkte Stöße ionisieren, da wegen der mit dem hohen Druck verbundenen großen Stoßhäufigkeit zwischen zwei Stößen meist weniger Energie aus dem elektrischen Feld aufgenommen wird, als zur Ionisierung erforderlich ist. Statt dessen heizen die im Katodenfallraum beschleunigten Elektronen durch elastische Stöße das Gas der Bogensäule stark auf und machen dadurch die thermische Ionisierung möglich. Diese ist für die meisten Bogenentladungen charakteristisch und die Ursache dafür, daß der Spannungsbedarf von Bogenentladungen in der Regel größenordnungsmäßig zehnmal kleiner ist als in vergleichbaren Glimmentladungen.
Ein Beispiel für eine frei brennende Bogenentladung ist der bekannte *Lichtbogen.* Eine Sonderform der Bogenentladung bildet der sogenannte *Gerdienbogen,* bei dem die Bogensäule vermittels eines Wasserwirbels stabilisiert und auf einen kleinen Durchmesser eingeschnürt wird. Dadurch können sich in ihr Plasmatemperaturen bis 50000 K, in neueren Anordnungen sogar bis 80000 K ausbilden. Das sind die höchsten bisher bei Atmosphärendruck über längere Zeit erzeugten Temperaturen. Sonderformen von Bogenentladungen stellen auch die elektrischen Funkenentladungen dar, die weiter nichts sind als sehr kurzzeitige und wenig definierte Bogenentladungen.

9.1.7.3.4. *Höchststromentladungen.* Dazu werden alle elektrischen Gasentladungen gezählt, bei denen die Entladungsstromdichte so außerordentlich groß ist, daß das mit dieser verbundene innere Eigenmagnetfeld einen bestimmten Einfluß auf den Entladungsablauf ausübt. Durch dieses starke Magnetfeld wird das Säulenplasma schlauchartig zusammengeschnürt. Man bezeichnet diese Erscheinung als magnetischen *Pincheffekt* und solche Höchststromentladungen als *Pinchentladungen.* Die radiale Kontraktion der positiven Säule hält jedoch nur solange an, bis der Druck des Magnetfelds durch den im Säulenkern durch die Kompression entstehenden hohen Druck der Plasmateilchen kompensiert bzw. überkompen-

siert wird. Es entsteht auf diese Weise eine Reihe aufeinanderfolgender Kontraktionen und Aufweitungen. Diese sind mit einem schwingenden Verlauf von Entladungsspannung und Entladungsstromstärke verbunden. Entladungen dieser Art spielen in der Hochtemperaturplasmaforschung eine wichtige Rolle, da die im Säulenplasma bei den Kontraktionen entstehenden Temperaturen beträchtlich sind.

9.1.7.3.5. *Sonderformen und Übergänge.* Ähnlich wie mit Gleichspannung können Gasentladungen auch mit Wechselspannung betrieben werden. Man bezeichnet sie dann als *Wechselstromentladungen.* Bei ihnen vertauschen Katoden- und Anodenfall im Rhythmus des Polaritätswechsels der Elektroden ständig ihre Rollen, während die positive Säule weitgehend unangetastet bleibt. Die in den Elektrodenfällen vorhandenen Raumladungen werden dabei ständig auf- und abgebaut. Bei höheren Frequenzen der Wechselspannung kann der Raumladungsaufbau und -abbau den Feldwechseln nicht mehr folgen, so daß sich dann für die Entladung andere als die beschriebenen Eigenschaften ergeben; z. B. spielen für solche *Hochfrequenzentladungen* die Sekundärprozesse an der Katode bzw. an den Elektroden keine entscheidende Rolle mehr. Deshalb kann bei Hochfrequenzentladungen auch gänzlich auf innere Elektroden im Entladungsgefäß verzichtet werden. Die Anregung und Ionisierung des im Entladungsgefäß eingeschlossenen Gases erfolgt dann durch kapazitive oder induktive Ströme von außen her. Man braucht dazu das elektrodenlose Entladungsgefäß nur in ein ausreichend starkes Hochfrequenzfeld eines Kondensators oder einer Spule zu bringen (Abb. 9.1.7-5).

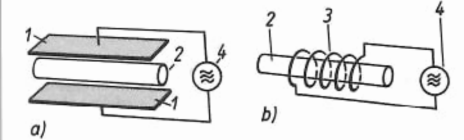


Abb. 9.1.7-5 Plasmaerzeugung in Hochfrequenzentladungen; *1* Platten eines Hochfrequenzkondensators, *2* elektrodenloses (gasgefülltes) Entladungsrohr, *3* Windungen einer Hochfrequenzspule, *4* leistungsstarker Hochfrequenzgenerator

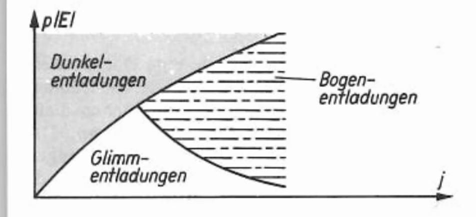


Abb. 9.1.7-6 Existenzbereiche von Dunkel-, Glimm- und Bogenentladungen in Abhängigkeit von Stromdichte j, Gasdruck p und elektrischer Feldstärke $|\boldsymbol{E}|$

Wechselstrom- und Höchststromentladungen sind zeitveränderliche Entladungen und können daher mit den stationären oder zeitunveränderlichen Entladungsformen der Dunkel-, Glimm- und Bogenentladung nicht ohne weiteres verglichen werden. Andererseits ist es möglich, die stationären Entladungsformen der Dunkel-, Glimm- und Bogenentladung unter günstigen Verhältnissen durch Änderung der Entladungsstromdichte direkt ineinander überzuführen. Ob solche Entladungen dann als Dunkel-, Glimm- oder Bogenentladungen existieren, hängt hauptsächlich vom Gasdruck, von der im Katodenfallraum erzeugten elektrischen Feldstärke und von der Stromdichte der Entladung ab (Abb. 9.1.7-6).

9.1.8. Technische Anwendung des Plasmazustands

Zu diesen Anwendungen können, strenggenommen, nur die gerechnet werden, bei denen der Plasmazustand in seiner charakteristischen Form als Medium aus Elektronen, Ionen und Neutralteilchen auftritt. Elektronen- und Ionengeräte, die mit Elektronen oder Ionen allein arbeiten, gehören nicht dazu.

9.1.8.1. Anwendungen in der Lichttechnik. Die vom Plasma ausgesandte elektromagnetische Strahlung im sichtbaren und ultravioletten Spektralbereich wird seit vielen Jahrzehnten zur Lichterzeugung ausgenutzt. Solche *Gasentladungslichtquellen* zeichnen sich durch große Lichtmenge und große Leuchtdichte aus. Im Unterschied zur normalen Glühlampe, bei der ein großer Teil der zugeführten elektrischen Energie als unsichtbare und damit unerwünschte Wärmestrahlung auftritt, wird die den Gasentladungslichtquellen zugeführte elektrische Energie weitgehend in Lichtenergie umgewandelt; z. B. beträgt der Wirkungsgrad moderner Gasentladungslichtquellen 80 bis 90%.

Eine der ältesten und heute nur noch in älteren Projektionsapparaten und in großen Scheinwerfern gebrauchte Gasentladungslichtquelle ist der *Kohlelichtbogen.* Er zeichnet sich durch besonders große Lichtmenge und große Lichtdichte aus. Allerdings entstammt die von ihm abgegebene Strahlung weniger dem Plasma, sondern mehr der durch die Plasmaelektronen auf helle Weißglut erhitzten Bogenanode, die durch ihren trichterförmigen Abbrand den sogenannten *positiven Krater* ausbildet (Abb. 9.1.8-1). In so ziemlich allen anderen Gasentladungslichtquellen hingegen wird die Strahlung

des Plasmas der positiven Säule zur Lichterzeugung ausgenutzt. Hierzu gehören die verschiedensten Typen der *Glimmentladungslichtquellen.* Das sind Gasentladungsstrecken, die unter vermindertem Druck mit Gasen, wie Stickstoff, Kohlendioxid, Natriumdampf, oder auch mit Mischungen von mehreren Gasen oder Dämpfen gefüllt sind und für alle möglichen

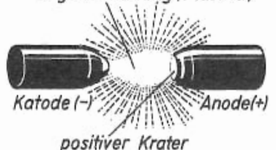


Abb. 9.1.8-1 Kohlelichtbogen

Beleuchtungszwecke Anwendung gefunden haben. Hierher gehören auch die bekannten Signal- und Kontrollampen, in denen vielfach statt der Strahlung der positiven Säule die des negativen Glimmlichts ausgenutzt wird. Durch Füllung mit Edelgasen oder deren Mischungen lassen sich mit ihnen die verschiedensten Signalfarben erzielen. Andere Gasentladungslichtquellen nutzen die starke Strahlung der positiven Säule von Bogenentladungen aus, z. B. alle Formen der *Hoch-* und *Höchstdrucklampen.* Das sind Gasentladungslichtquellen, bei denen infolge des hohen Drucks bis zu 10 MPa die positive Säule stark eingeschnürt ist bzw. sich bei kleinem Elektrodenabstand quasi-punktförmig ausbildet. Sie sind meist mit Quecksilberdampf oder Edelgasen gefüllt. Wegen des hohen Drucks ist ihre Ausstrahlung über weite Gebiete des Spektrums fast kontinuierlich. Solche quecksilberdampfgefüllte Hoch- und Höchstdrucklampen finden z. B. für medizinische Zwecke als Höhensonne Verwendung. Entsprechende natriumdampfgefüllte Lichtquellen werden wegen ihres kontrastreichen gelben Lichts gern als Straßenbeleuchtung verwendet. Xenonhöchstdrucklampen schließlich weisen ein dem Sonnenlicht sehr ähnliches Spektrum auf und sind daher als *Tageslichtlampen* für die farbgetreue Farbfilmwiedergabe unentbehrlich geworden. Vielfach sind die Gasentladungslichtquellen auf ihrer inneren Gefäßwand mit einem Leuchtstoffbelag, einem *Luminophor*, versehen, der die Aufgabe hat, den hohen Anteil an ultravioletter Strahlung in sichtbare Strahlung umzuwandeln und somit die Ausbeute an sichtbarem Licht weiter zu erhöhen. Die zunehmende Substitution von Glühlampen durch Leuchtstoffröhren sowie die Erhöhung der Lichtausbeute mittels verbesserter Leuchtstoffe (↑ 8.8.3.3.5.) bilden wichtige Maßnahmen zur Einsparung von Elektroenergie.

Weiterhin sind noch die wohl jedem bekannten *Reklameleuchtröhren* zu nennen. Sie müssen wegen der oftmals notwendigen großen Länge der strahlenden positiven Säule mit ziemlich hoher Spannung betrieben werden und werden daher auch als *Hochspannungsleuchtröhren* bezeichnet. Sie sind gewöhnlich mit Edelgasen oder deren Mischungen und gegebenenfalls noch mit Metalldampfzusätzen gefüllt, wodurch ihr Licht über eine weite Farbpalette variiert werden kann. Für wissenschaftliche und technische Zwecke sind schließlich viele Typen von *Spektrallampen* entwickelt worden, die gewöhnlich mit Reinstgasen gefüllt sind und durch ihr definiertes Linienspektrum in der chemischen und technischen Spektralanalyse zu Eich- und Vergleichszwecken dienen. Zu ihren ältesten Typen gehören die oft in den Sammlungen von Schulen erhalten gebliebenen skurril geformten Geißlerröhren.

9.1.8.2. Anwendungen in der Elektrotechnik. Eine der ältesten Anwendungen des Plasmazustands in der Elektrotechnik findet sich bei der *Gleichrichtung* starker Wechselströme. Die dazu benutzten Gasentladungseinrichtungen sind gewöhnlich mit Quecksilberdampf gefüllt und werden daher als *Quecksilberdampfgleichrichter* bezeichnet (Abb. 9.1.8-2). Dazu werden Glas-, bei größeren Ausführungen auch Metallgefäße, verwendet. Die Katode besteht meist aus flüssigem Quecksilber, während die Anoden ziemlich großflächig und aus Eisen oder Graphit sind. In diesen Entladungsgefäßen brennt infolge der angelegten Wechselspannung

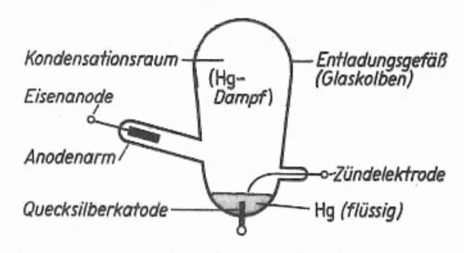


Abb. 9.1.8-2 Prinzipieller Aufbau eines Quecksilberdampfgleichrichters

zwischen Katode und Anode eine Quecksilberdampfbogenentladung, durch deren Säulenplasma wegen seiner guten elektrischen Leitfähigkeit ein hoher Strom fließen kann. Die Gleichrichterwirkung solcher Anordnungen beruht darauf, daß *nur* die Quecksilberelektrode die für den Stromtransport notwendigen Elektronen liefern kann, so daß nur dann ein Strom fließen kann, wenn bei angelegter Wechselspannung die Quecksilberelektrode Katode ist; im umgekehrten Fall bleibt die Entladungsstrecke für den Stromtransport gesperrt.

Kleinere Gasentladungsgleichrichter haben vielfach an Stelle des Quecksilberdampfs eine Edelgasfüllung. In diesem Fall muß die Elektronenemission durch eine elektrische Heizung der Katode bewerkstelligt werden. Gasentladungsgleichrichter dieser Art werden als gasgefüllte *Glühkatodengleichrichter* bezeichnet. Sie können wegen der Ionisierung der Gasfüllung durch die aus der Katode emittierten Elektronen ebenfalls größere Ströme gleichrichten. Eine weitere Sonderform stellt der *Glimmentladungsgleichrichter* dar, der nur zur Gleichrichtung schwächerer Ströme geeignet ist. Seine Gleichrichterwirkung beruht auf der sehr verschiedenen Größe, die seine beiden Entladungselektroden haben. Auch die bekannten *Stabilisatorröhren*, die in Stromversorgungsgeräten vielfach zur Spannungsstabilisierung verwendet werden, sind Gasentladungseinrichtungen. In ihnen nutzt man die besondere Eigenschaft der normalen Glimmentladung, einen recht konstanten Katodenfall zu haben, zur Stabilisierung aus.

Eine weitere wichtige Anwendung hat der Plasmazustand in der Elektrotechnik als *Schaltmittel* gefunden. Während der zwischen den Kontakten eines mechanischen Schalters gewöhnlich entstehende kurzzeitige Lichtbogen, der *Schaltlichtbogen*, die Schaltzeit verlängert, auf die Schalterkontakte zerstörend wirkt und somit sehr unerwünscht ist, nutzt man in den Gasentladungsschaltern die besonderen Eigenschaften des Plasmas geradezu zum sachgemäßen Schalten großer Ströme aus. Durch Einbau einer dritten Elektrode zwischen Katode und Anode hat man es bei Gasentladungsstrecken weitgehend in der Hand, gesteuerte Schalthandlungen auszuführen. Ein solcher einfacher *Gasentladungsschalter* ist die mit einer dritten Elektrode, einer Zünd- oder Triggerelektrode, ausgerüstete Funkenstrecke, die damit zur *Schaltfunkenstrecke* wird. Ganz ähnlich können auch gasgefüllte Entladungsröhren als Schaltmittel verwendet werden. Sie sind gewöhnlich mit einem Metallgitter oder einem Zündstift aus Halbleitermaterial als dritter Elektrode ausgerüstet und werden als *Thyratrons* bzw. *Ignitrons* bezeichnet.

Außer in der Starkstromtechnik haben Schalteinrichtungen auf Plasmagrundlage auch sehr breite Verwendung in der Radartechnik gefunden. Da hier zum Senden und Empfangen der Signale gewöhnlich die gleiche Antenne verwendet wird, müssen zur einwandfreien Trennung von Sende- und Empfangsignalen außerordentlich schnell arbeitende Umschalter benutzt werden. Spezielle *Gasentladungsschaltröhren* erfüllen diese Forderung, da sie wegen der kurzen Rekombinationszeiten des Plasmas mit Leichtigkeit 10^5 bis 10^6 einzelne Schalthandlungen je Sekunde ausführen können.

9.1.8.3. Anwendungen in der Metallurgie und Metallbearbeitung. Hierbei nutzt man vor allem die mit dem Plasmazustand verbundenen außerordentlich hohen Temperaturen aus; z. B. wird bereits seit längerer Zeit die hohe Temperatur der positiven Säule des in atmosphärischer Luft brennenden Lichtbogen zum *Schweißen* verwendet. Bei den *Schweißlichtbögen* brennt die Entladung entweder zwischen zwei Elektroden, die in unmittelbare Nähe der Schweißstelle gebracht werden, oder zwischen einer Elektrode und dem Schweißgut selbst. In jedem Fall reichen die z. B. mit Kohleelektroden erzeugbaren Temperaturen von 5000 bis über 6000 K aus, um alle bekannten Metalle zum Schmelzen zu bringen. Vielfach wird in den Schweißlichtbogen, der hierbei zwischen Wolframelektroden brennt, zusätzlich molekulares Wasserstoffgas eingeblasen. Dieses dissoziiert im Bogen infolge der hohen Temperatur, um nach Wiederverlassen des Bogens wieder zu molekularem Wasserstoff zu rekombinieren. Dabei wird die im Bogen zur Dissoziation aufgewendete Energie wieder frei und beim Schweißvorgang zusätzlich an der interessierenden Stelle zum Metallschmelzen ausgenutzt. Dieses besonders wirkungsvolle Schweißverfahren wird als *Arcatomverfahren* bezeichnet.

Auch zum Schmelzen von Metallen und ähnlichen hochtemperaturfesten Stoffen hat der elektrische Lichtbogen breite Anwendung gefunden; z. B. können in den *Lichtbogenöfen*, in denen zwischen Graphitelektroden und dem Schmelzgut Ströme von tausend und mehr Ampere fließen, die meisten Metalle erschmolzen werden.

In neuerer Zeit werden Schweiß-, Schneid- und Schmelzarbeiten statt im Lichtbogen vielfach mit einem *Plasmastrahl* ausgeführt. Dieser wird dadurch erzeugt, daß man mittels eines schnellen Gasstroms das Plasma einer stromstarken Bogenentladung durch eine in der Bogenanode angebrachte düsenförmige Öffnung nach außen bläst. Anordnungen dieser Art werden als

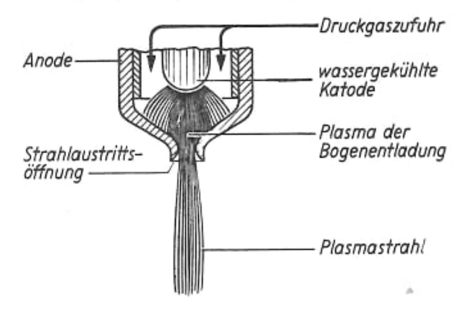


Abb. 9.1.8-3 Prinzipieller Aufbau des Plasmabrenners

Plasmabrenner bezeichnet (Abb. 9.1.8-3). Der sich außerhalb der Bogenentladung ausbildende Plasmastrahl ist das eigentliche Werkzeug. Des weiteren können mit dem Plasmastrahl durch Zugabe pulverförmiger Metallkarbide auch hochtemperatur- und verschleißfeste Überzüge auf Werkstücke aufgesprüht werden. Diese Technologie wird als *Plasmaspritzen* bezeichnet. Schließlich nutzt man in wachsendem Maße die bei ständiger Einwirkung eines Funkens auf Metallteile entstehende Metallabtragung in der Metallverarbeitung aus. Dieser als *Funkenerosion* bezeichnete Vorgang ist insbesondere zur Feinbearbeitung sehr kleiner Werkstücke geeignet.

9.1.8.4. Anwendungen in der Chemie. Außerordentlich merkwürdig ist im Plasmazustand das chemische Verhalten der Elemente. Da dies durch die Anordnung und die Anzahl der Elektronen in ihrer äußeren Elektronenschale bestimmt wird, hängt es in starkem Maße davon ab, ob sich ein Element im ionisierten Zustand befindet, ob es ein oder mehrere Elektronen aus seiner Elektronenhülle verloren hat oder nicht. Während z. B. Edelgase bei normalen Temperaturen nicht miteinander oder mit anderen Stoffen reagieren können, weil sie über eine abgeschlossene äußere Elektronenschale verfügen, werden sie im ionisierten Zustand chemisch äußerst aktiv; z. B. sind in Edelgasplasmen schon oft Verbindungen von Edelgasen untereinander und mit anderen Stoffen beobachtet worden. Ähnlich zeigen auch alle anderen Elemente im Plasmazustand ganz andere Wertigkeiten und Affinitäten als im klassischen Temperaturbereich. Dadurch tritt im Plasmazustand eine Fülle neuer, bei niedrigeren Temperaturen nicht beobachtbarer chemischer Verbindungen auf. Mit der Erforschung und Nutzung dieser neuen chemischen Verbindungen im Plasmazustand beschäftigt sich neuerdings eine eigene Forschungsrichtung, die *Plasmachemie*.

Aber auch in der klassischen Chemie hat der Plasmazustand Anwendungen gefunden, z. B. bei dem allgemein bekannten Verfahren der Salpetersäureherstellung nach BIRKELAND-EYDE. Bei diesem wird das notwendige Zwischenprodukt Stickstoffmonoxid, das in der Natur gewöhnlich nicht vorkommt, *plasmachemisch* aus der Luft erzeugt. Technisch wird dabei ein starker Luftstrom durch einen durch ein Magnetfeld aufgeweiteten Wechselstromlichtbogen geblasen; dabei bildet sich in seiner heißen Zone aus dem Stickstoff und dem Sauerstoff der Luft Stickstoffmonoxid, das beim Wiederverlassen des heißen Lichtbogens als stabiles Reaktionsprodukt erhalten bleibt.

Eine andere, weniger bekannte Anwendung des Plasmazustands in der Chemie findet sich bei der künstlichen *Diamantherstellung*. Die bei dieser durch Umkristallisation von Graphit im heißen Plasma bei höchsten Drücken entstehenden Diamantkristalle sind zwar klein, aber durchaus vollwertig. (↑ 9.3.9.1.).

9.1.8.5. Anwendung als Motor und Dynamomaschine. Wird durch einen metallischen Leiter zwischen zwei Magnetpolen quer zu den magnetischen Kraftlinien ein elektrischer Strom geschickt, so tritt an dem Leiter eine Kraft auf, die ihn quer zum Magnetfeld und quer zur Stromrichtung in Bewegung setzt. Diese Erscheinung liegt unter anderem der Umwandlung von elektrischer in mechanische Energie in allen Elektromotoren zugrunde. Genau derselbe Vorgang findet statt, wenn der Strom statt durch einen metallischen Leiter durch ein zwischen Magnetpolen gelagertes Plasma fließt. Auch hier tritt infolge der Wechselwirkung zwischen dem im Plasma durch den Stromfluß erzeugten inneren Magnetfeld und dem von außen wirkenden Magnetfeld eine Beschleunigung des Plasmas quer zur Stromrichtung und zum äußeren Magnetfeld auf. Auf dieser Grundlage arbeiten in der Forschung zahlreiche Verfahren, mit denen Plasmen auf sehr hohe Geschwindigkeiten beschleunigt werden können. Eine ältere Anwendung dieses Prinzips stellen die *elektromagnetischen Konduktionspumpen* dar, die allerdings statt zur Bewegung von Plasmen meist zur Förderung flüssiger Metalle Verwendung finden. Eine ausgesprochen plasmaphysikalische Anwendung dieses Prinzips findet sich in den sogenannten *Plasmakanonen*, die zur impulsmäßigen Beschleunigung von Plasmen benutzt werden. Die mit ihnen beschleunigten Plasmapakete können Geschwindigkeiten von vielen 100 kms^{-1} erreichen. Einrichtungen dieser Art dürften in Zukunft für den Antrieb von Raketen oder Raumschiffen größere Bedeutung erlangen.

Während in den aufgeführten Anordnungen die dem Plasma vermittels des Stroms zugeführte elektrische Energie bei der Beschleunigung des Plasmas in Bewegungsenergie umgewandelt wird, kann andererseits die dem Plasma innewohnende thermische Energie in elektrische Energie umgeformt werden. Die Wirkung dieser *magnetohydrodynamischen* oder *MHD-Generatoren* (Abb. 9.1.8-4) beruht auf der Umkehrung des oben beschriebenen Vorgangs. Ein beim Herausströmen aus einer Brennkammer zwischen zwei Magnetpolen hindurchfließendes Plasma vermag zwischen zwei an geeigneten Stellen eingeführten Elektroden eine elektrische Spannung aufzubauen. Verbindet man diese Elektroden über einen in einem äußeren Stromkreis befindlichen Verbraucher, so wird dem strömenden Plasma Energie durch Abbremsen entzogen und diese dem Verbraucher zugeführt. Da hierbei die Umwandlung von thermischer in elektrische

Energie direkt erfolgt, haben solche Anordnungen einen höheren Wirkungsgrad als die herkömmlichen Verfahren, als die mit Dampfmaschinen oder Dampfturbinen angetriebenen Dynamomaschinen. Wegen ihres höheren Wirkungsgrads dürften solche magnetohydrodynamischen Generatoren in Zukunft zur besseren Energieausnutzung in Wärme- und Kernkraftwerken große Bedeutung erlangen.

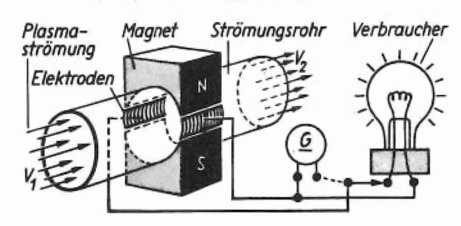


Abb. 9.1.8-4 Schematische Darstellung eines MHD-Generators; *G* Generator, V_1, V_2 Plasmaströmungsgeschwindigkeit vor und hinter der Anordnung

9.1.8.6. Anwendung zur Energieerzeugung durch gesteuerte Kernfusion. Die wichtigste Triebkraft bei der weiteren Entwicklung der Plasmaphysik bildet gegenwärtig und künftig das Bemühen zur Erschließung der gesteuerten Kernfusion für die Energieerzeugung. Große Forschungsprogramme in der UdSSR, den USA u. a. Ländern sind darauf ausgerichtet, die für eine Verschmelzung leichter Atomkerne (H, D, Li) notwendigen, extrem hohen Plasmatemperaturen und -dichten zu erzeugen und über ausreichend lange Zeitintervalle aufrechtzuerhalten. Die damit verbundenen Probleme werden am Ende dieses Kapitels behandelt (↑ 9.4.).

9.2. Erzeugung und Messung von Vakuum

9.2.1. Übersicht

Obwohl das ideale Vakuum als leerer Raum nicht zu den Stoffzuständen gerechnet werden kann, sollen im folgenden einige Ausführungen zur Erzeugung und Messung von Vakua eingefügt werden. Der Zusammenhang mit dem Gegenstand eines Werks über die Struktur der Materie und speziell die Verbindung mit extremen Stoffzuständen leitet sich aus einer Reihe von Wechselbeziehungen ab.

Ein stoffleerer Raum, also das ideale Vakuum, ist praktisch nicht realisierbar. Selbst im Ultrahochvakuum der besten irdischen Apparaturen sind noch Restgase mit einem Druck von 10^{-8} bis 10^{-14} Pa vorhanden. Das entspricht etwa 10^5 bis 10^{11} Atomen je Kubikzentimeter. Die realisierbaren Vakua stellen also Gase in extremer Verdünnung dar.

Auch im *interstellaren Raum* (↑ 10.3.) findet sich noch stoffliche Materie in einer Konzentration von etwa einem Atom je Kubikzentimeter, und innerhalb des Planetensystems ist die Materiedichte sogar um Größenordnungen höher. Kurioserweise erscheinen in letzter Zeit zunehmend Publikationen, die sich mit dem Problem der Erzeugung hoher Vakua in Weltraumstationen befassen. Durch verdampftes Material der Satelliten selbst ist das Vakuum in ihrer Umgebung nämlich nur mäßig, und man wendet beispielsweise Schutzschilde an, um in ihrem Schatten gute Vakuumbedingungen aufrechtzuerhalten.

Die Erforschung der Struktur der Materie ist untrennbar mit dem Einsatz immer leistungsfähiger Vakuumanlagen verbunden. Erst nach weitgehender Entfernung der Luft können sich Teilchenstrahlen von Elektronen, Ionen oder Neutralatomen einigermaßen ungestört ausbreiten. In Elektronenmikroskopen, Teilchenbeschleunigern oder vielen Geräten zur Erforschung der Struktur und Zusammensetzung von Stoffen muß daher ein möglichst hohes Vakuum aufrechterhalten werden.

Vor allem in der Festkörper- und Halbleiterphysik bildet das Arbeiten im Vakuum oder unter Schutzgas eine wichtige Voraussetzung, um Verunreinigungen auszuschließen. Besonders extreme Anforderungen werden in dieser Hinsicht bei Untersuchungen zu Vorgängen an Oberflächen gestellt, und hierfür wird vorzugsweise extremes Ultrahochvakuum eingesetzt.

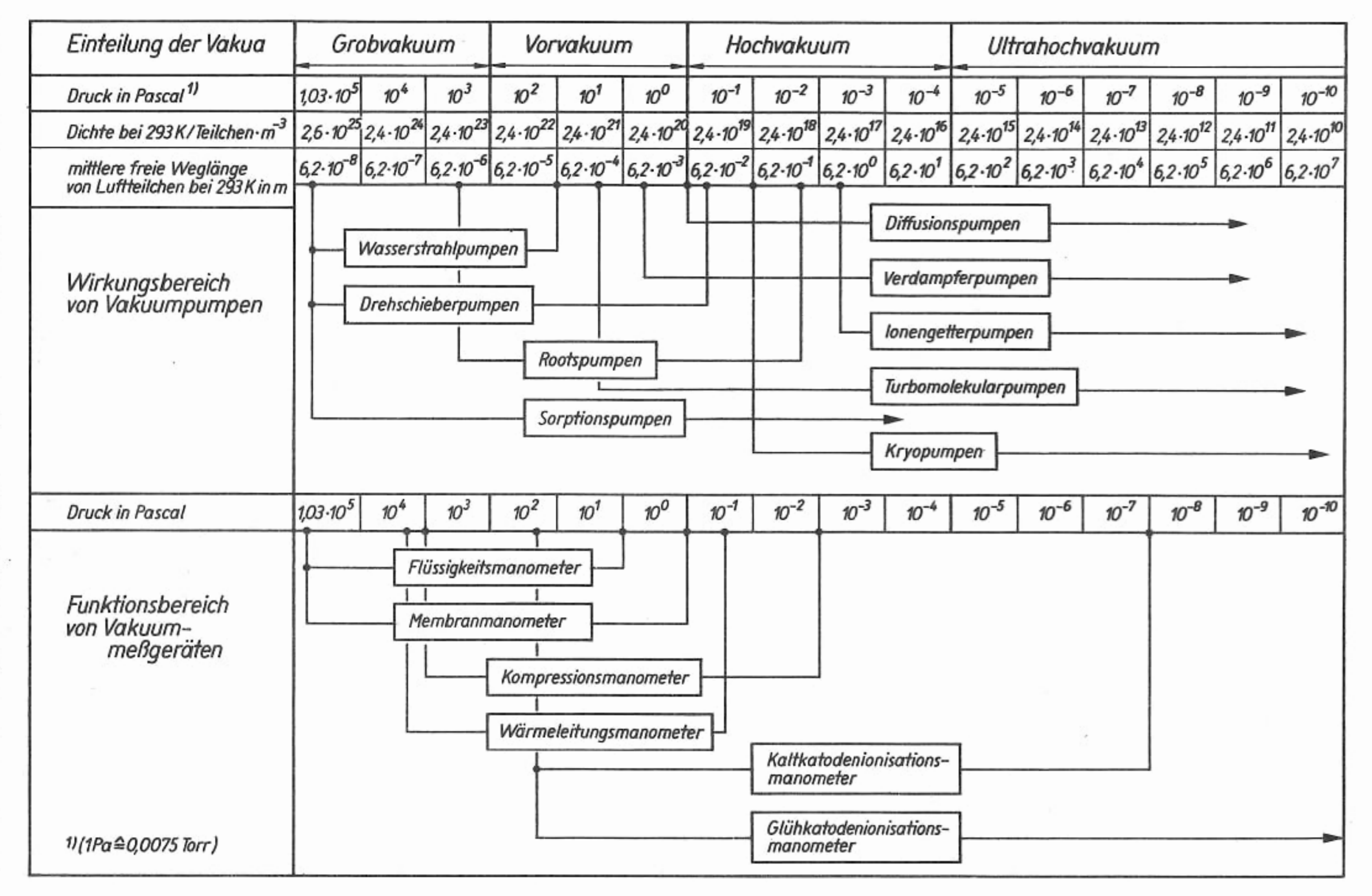
Schließlich werden auch viele Prozesse zur Herstellung und Veredlung von Werkstoffen im Vakuum durchgeführt. Das gilt beispielsweise für wichtige Beschichtungsverfahren, für die Ionenimplantation oder für die Herstellung sehr reiner Einkristalle. Vakuumverfahren erlangen auch im technischen Maßstab eine steil ansteigende Bedeutung.

Nur hinzuweisen ist auf das im Rahmen der Elementarteilchenphysik (↑ 3.1.2.) herangereifte Problem, ob das ideale Vakuum wirklich leer ist. Gewisse Anzeichen sprechen dafür, daß das Vakuum polarisierbar ist und demnach eine physikalische Struktur haben muß.

In Tabelle 9.2.1-1 wird eine Übersicht über die Charakterisierung der für Forschung und Technik wichtigen Vakuumbereiche gegeben. Die Zusammenstellung vermittelt zugleich Informationen bezüglich der Pumpen und Druckmeßgeräte, die in den einzelnen Bereichen zur Anwendung kommen.

9.2.2. Wirkprinzip einiger Vakuumpumpen

Bei der *Drehschieberpumpe* (Abb. 9.2.2-1a) läuft im zylindrischen Pumpgehäuse der exzentrisch gelagerte Rotor mit zwei oder



Tab. 9.2.1-1 Die wichtigsten Vakuumbereiche, Wirkungsbereiche von Vakuumpumpen sowie Funktionsbereiche von Vakuummeßgeräten

mehreren Drehschiebern, die durch die Feder- und Fliehkraft auseinandergedrückt werden und entlang der Gehäusewand gleiten. Beim Drehen des Rotors wird der Schöpfraum nach Erreichen des maximalen Volumens durch den nachfolgenden Drehschieber abgeschlossen, danach erfolgt eine Verdichtung des eingeschlossenen Gases und der Ausstoß durch das sich öffnende Auslaßventil.

Im Gehäuse der *Rootspumpe* (Abb. 9.2.2-1 b) sind zwei Rotoren mit achtförmigem Querschnitt gelagert, die sich ohne Reibung untereinander und mit der Gehäusewand gegenläufig drehen. Dabei wird ein bestimmtes Gasvolumen durch den Saugstutzen angesaugt und zwischen dem Rotor und der Gehäusewand zum Druckstutzen transportiert.

In der dargestellten dreistufigen fraktionierenden *Öldiffusionspumpe* (Abb. 9.2.2-1 c) wird der im Siedegefäß erzeugte und durch die drei ineinander angeordneten Steigrohre zu den Umkehrdüsen geleitete Öldampf auf hohe Geschwindigkeit beschleunigt und breitet sich im Pumpraum aus. Hier gelangt das Gas in den Dampfstrahl, wird verdichtet und von der Hoch- zur Vorvakuumseite transportiert, wo es mit Hilfe einer Vorvakuumpumpe abgesaugt wird. Der Öldampf kondensiert an der äußeren gekühlten Gehäusewand und läuft in das Siedegefäß zurück.

Bei den *Getterpumpen* wird das Gas vorwiegend durch physikalische oder chemische Adsorption an einer mit dem Gettermaterial belegten Festkörperoberfläche oder durch Verbindungsbildung gebunden. In der *Verdampferpumpe* wird das Gettermaterial, z. B. Titanium, verdampft, und auf einer Festkörperoberfläche abgeschieden.

Bei der *Ionenzerstäuberpumpe* (Abb. 9.2.2-1 d) wird durch Anlegen einer Beschleunigungsspannung U_R eine magnetfeldunterstützte Gasentladung gezündet. Die Oberfläche der vorwiegend aus Titanium bestehenden Katode wird durch die auftreffenden positiven Ionen zerstäubt. Das abgestäubte Gettermaterial schlägt sich insbesondere auf den Anodenflächen nieder. Aufgrund unterschiedlicher Getterwirkungen ist die Sauggeschwindigkeit der Getterpumpen für aktive Gase wesentlich höher als für Edelgase.

Die Pumpwirkung von *Turbomolekularpumpen* (Abb. 9.2.2-1 e) beruht auf dem Effekt der Reibung eines Gases, insbesondere auf der Impulsübertragung an schnell bewegten Wänden ähnlich dem Turbinenprinzip. Die Restgasteilchen werden zwischen dem Rotor (Drehzahl etwa 20000 Umin^{-1}) und dem Stator vom Saugstutzen zum Vorvakuumstutzen transportiert.

Die Tatsache, daß bei hinreichend niedrigen

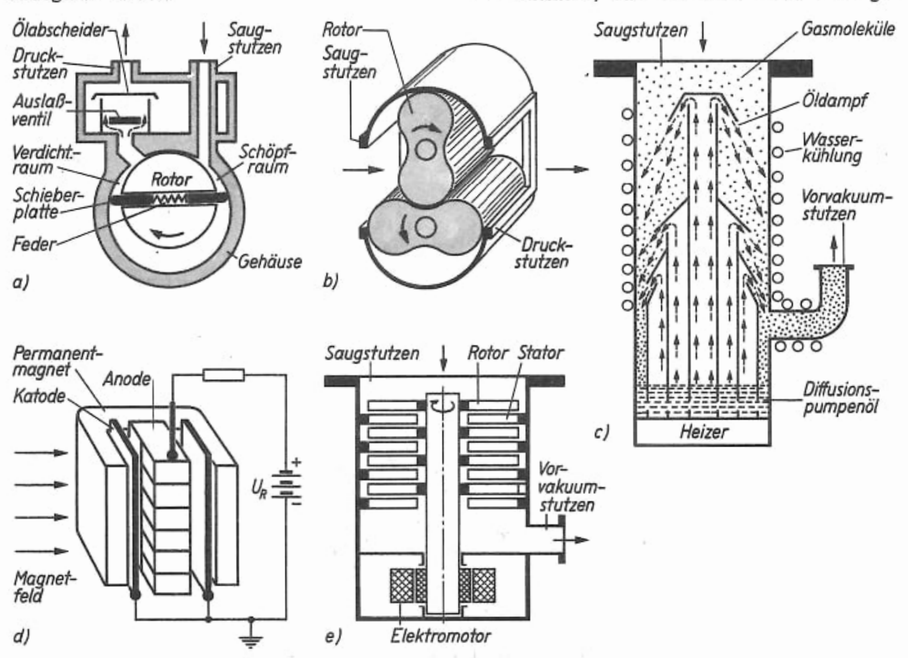


Abb. 9.2.2-1 a) Drehschieberpumpe, b) Rootspumpe, c) Öldiffusionspumpe d) Ionenzerstäuberpumpe, e) Turbomolekularpumpe

Temperaturen der Gleichgewichtsdruck über der festen Phase eines Gases oder über einem Gasadsorptionsmittel sehr niedrig ist, wird bei den *Kryopumpen* zur Vakuumerzeugung genutzt. Bei den *Kondensationspumpen* wird das Gas an tiefgekühlten Flächen „ausgefroren", bei der *Adsorptionspumpe* hingegen erfolgt eine Adsorption des Gases durch Adsorptionsmittel. Um Ultrahochvakuum zu erzeugen, müssen die Kühlflächen mit Hilfe von flüssigem Helium auf etwa 4 K gehalten werden. Neuere Kryopumpen sind mit kompakten Kühlaggregaten ausgestattet, in denen das Helium im Kreislauf genutzt wird.

9.2.3. Wirkprinzip einiger Vakuummeßgeräte

Zur absoluten Messung geringer Drücke eignet sich das *Kompressionsmanometer* nach Mc Leod (Abb. 9.2.3-1a), dessen Funktion auf der Anwendung des Boyleschen Gesetzes (↑ 6.2.2.) beruht. Durch Anheben der Quecksilbersäule wird das Meßvolumen von der Zuleitung zum Vakuumgefäß abgetrennt und bis auf das Kapillarvolumen komprimiert. Der Druck im Vakuumgefäß ist dem Höhenunterschied der beiden Quecksilbermenisken direkt proportional.

Bei den *Wärmeleitungsmanometern* nutzt man die Druckabhängigkeit der Wärmeableitung von einem geheizten Draht, d. h. die Heizdrahttemperatur, zur Druckmessung aus. Im *Piranimanometer* wird die Temperatur des Heizdrahts über die Temperaturabhängigkeit seines elektrischen Widerstands ermittelt; im *thermoelektrischen Vakkummeter* (Abb. 9.2.3-1b) hingegen dient ein Thermoelement zur Temperaturmessung des Heizdrahts. Da die Wärmeableitung und somit die Druckanzeige dieser Geräte von der Gasart abhängt, werden sie gegen ein Vakuummeter mit absoluter Druckanzeige eingemessen. Beim *Kaltkatodenionisationsmanometer* (Abb. 9.2.3-1c) erfolgt die Ionisierung von Restgasteilchen mit Hilfe von Elektronen, die durch Ionenbeschuß an der Katode ausgelöst wurden oder beim Ionisationsprozeß im Volumen entstehen und im angelegten elektrischen und magnetischen Feld beschleunigt und geführt werden. Hält man die Entladungsspannung und die Magnetfeldstärke konstant, dann ist der zwischen Katode und Anode fließende Entladungsstrom ein Maß für den Druck in der Vakuumkammer.

Im *Glühkatodenionisationsmanometer* (Abb. 9.2.3-1d) werden die Elektronen durch thermische Emission einer Glühkatode erzeugt und zwischen Katode und Anode auf die zur Ionisation der Restgasteilchen notwendige Energie beschleunigt. Die erzeugten Gasionen werden von einem negativ vorgespannten Ionenkollektor abgesaugt; der Kollektorstrom ist ein Maß für den Druck in der Vakuumkammer. Für den Bereich des extremen Ultrahochvakuums (UHV) wurden Ionisationsmanometer mit veränderter geometrischer Gestaltung entwickelt, sogenannte *Bayard-Alpert-Röhren.*

9.3. Hohe Drücke

9.3.1. Grundbegriffe

Zur Abgrenzung der Hochdruckphysik und -technik zu niederen Drücken gibt es keine allgemein akzeptierten Definitionen. An dieser Stelle soll zur Einschränkung der Thematik nur der Druckbereich über 100 MPa betrachtet werden.

Der Begriff des *hydrostatischen Drucks* ist im engeren Sinne nur für Flüssigkeiten und Gase

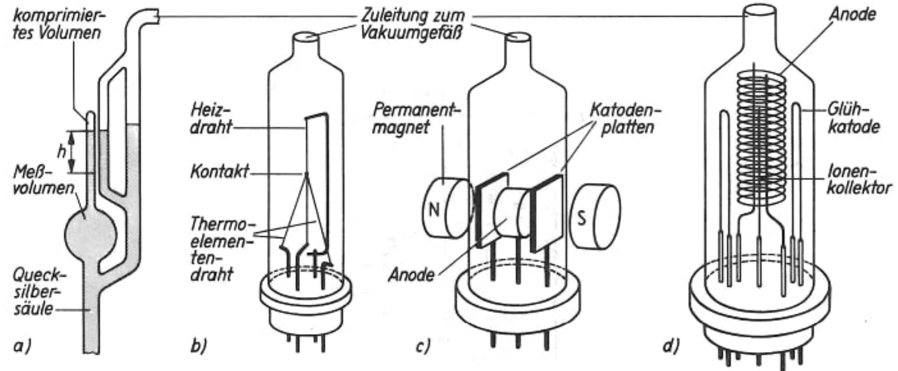


Abb. 9.2.3-1 a) Kompressionsmanometer (h Höhenunterschied der Quecksilbermenisken) b) thermoelektrisches Vakuummeter, c) Kaltkatodenionisationsmanometer, d) Glühkatodenionisationsmanometer

definiert. Wird ein derartiges deformierbares Medium langsam komprimiert, so ist die auftretende Kraftwirkung auf eine Bezugsfläche in dem Medium unabhängig von dem Ort und der Richtung der Fläche. Man definiert dann als Druck den Quotienten aus dem Betrag der Kraft und dem Flächeninhalt der Bezugsfläche. Wird dagegen ein fester Körper deformiert, so treten in ihm mechanische Spannungen auf, die im allgemeinen Fall orts- und richtungsabhängig sind. Sie können für einen Punkt durch die Angabe des Spannungstensors beschrieben werden. Der hydrostatische Fall ist als Sonderfall in dieser Beschreibung enthalten. Er tritt auf, wenn in der Hauptachsendarstellung die drei Hauptspannungen gleich und ortsunabhängig sind.
Eine theoretisch nicht ganz klar fixierte Situation liegt bei den sogenannten *quasihydrostatischen Drücken* vor. Man spricht von einem quasihydrostatischen Druck, wenn ein festes Druckmedium über die Elastizitätsgrenze hinaus bis in den plastischen Bereich deformiert wird, so daß plastisches Fließen einsetzt und sich das Material flüssigkeitsähnlich verhält.
Bezüglich der Zeitabhängigkeit ist zwischen *statischen* und *dynamischen Drücken* zu unterscheiden. Ein statischer Druck liegt vor, wenn die zeitlichen Änderungen vernachlässigt werden können, genauer gesagt, wenn die Druckänderungen in Zeiträumen erfolgen, die groß gegen die Laufzeit einer elastischen Welle durch das betrachtete Volumen sind. Bei dynamischen Drücken kann die zeitliche Änderung nicht mehr vernachlässigt werden. In einem engeren Sinne betrachtet man als dynamische Drücke die Drücke beim Durchgang von *Stoßwellen* durch ein Medium, z. B. bei Explosionen oder beim Aufprall hochbeschleunigter Körper. Das vorliegende Kapitel soll sich weitgehend auf die statischen Drücke konzentrieren.

9.3.2. Prinzipien der Druckerzeugung

Bei den *Methoden der Druckerzeugung* ist aus technischen Gründen ein Unterschied zwischen hydrostatischen und quasihydrostatischen Drücken vorhanden.
Eine Aussage über die z. Z. maximal realisierbaren Druckwerte ist relativ nichtssagend, wenn man nicht gleichzeitig eine Angabe darüber macht, in welchem Volumen diese Werte erreicht worden sind. Die höchsten statischen Drücke, die bisher realisiert worden sind, liegen bei einigen 100 GPa in einem Volumen von wenigen Kubikmillimetern.
Eine genaue Angabe eines Rekordwerts ist nicht möglich, weil die *Druckmessung* in den Grenzbereichen mit sehr großen Unsicherheiten behaftet ist. Die maximal erreichbaren Druckwerte sind durch mehrere Materialparameter bestimmt. Im hydrostatischen Druckbereich ist eine Grenze bereits dadurch gegeben, daß die *Viskosität der Flüssigkeiten* mit steigendem Druck ansteigt, bis sie schließlich in den festen Zustand übergehen. Es gibt nur wenige Flüssigkeiten, die bei Raumtemperatur zur Druckerzeugung oberhalb 3 GPa einsetzbar sind, z. B. einige leichte Kohlenwasserstoffe, wie Kerosin oder Gemische von Pentan und Isopentan. Die höchsten Werte, etwas über 10 GPa, sind mit einem Gemisch von Methanol und Ethanol im Verhältnis von 4 : 1 erreicht worden. Die Benutzung von Druckübertragungsmitteln, die bei Zimmertemperatur gasförmig sind, führt zu schwierigen Dichtungsproblemen und außerdem ergeben sich erhebliche praktische Schwierigkeiten bei Gasen durch deren hohe Kompressibilität.

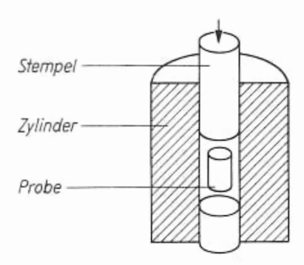


Abb. 9.3.2-1 Prinzip der Stempel-Zylinder-Hochdruckkammer

Bei der *quasihydrostatischen Druckerzeugung* ist eine Begrenzung durch die Festigkeit des Materials der Druckkammern gegeben, deshalb werden allgemein die Werkstoffe mit den besten Festigkeitseigenschaften eingesetzt, wie hochvergütete Stähle, Hartmetalle und in extremen Fällen kubisches Bornitrid und Diamant.

9.3.2.1. Druckerzeugung im hydrostatischen Bereich. Für industrielle Anwendungen stehen *Kompressoren* bis zu Drücken von etwa 0,5 GPa zur Verfügung. Mit ihrer Hilfe können relativ großvolumige Druckgefäße mit dem hochgespannten Druckmedium gefüllt werden. Es sind sowohl Kolben- wie Membrankompressoren in Gebrauch, in denen der Enddruck meist in mehreren Stufen erreicht wird.
Für die Erzeugung höherer hydrostatischer Drücke werden häufig *Stempel-Zylinder-Kammern* eingesetzt (Abb. 9.3.2-1). Sie bestehen im Prinzip aus einem dickwändigen Stahlrohr, das an beiden Enden durch Stahlstempel mit Dichtungen verschlossen ist. Das Rohr wird mit der als Druckmedium vorgesehenen Flüssigkeit gefüllt. Durch Einpressen der Stempel wird die Flüssigkeit komprimiert, wodurch in ihr der hohe Druck aufgebaut wird. Das Einpressen

erfolgt meist mit Hilfe von *hydraulischen Pressen* (Abb. 9.3.2-2). In der praktischen Ausführung dieses Prinzips wird häufig ein Stempel in einer festen Stellung im Rohr fixiert und nur der andere zum Komprimieren bewegt. Um eine hohe Festigkeit dieser Druckkammer und damit hohe Maximaldrücke zu erreichen, benutzt man sehr dickwandige Stahlrohre. Man

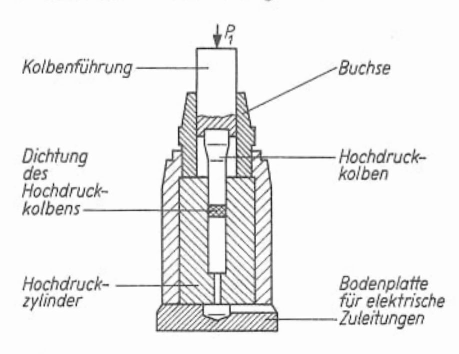


Abb. 9.3.2-2 Druckübersetzer in hydraulischen Pressen

kann jedoch die Festigkeit durch eine einfache Vergrößerung der Wandstärke nicht beliebig erhöhen. Diese Maßnahme wird immer weniger effektiv, weil das Material an der Außenwand des Rohrs bei sehr dicken Wänden schließlich kaum noch einen Einfluß auf die entscheidende Festigkeit an der Innenwand hat. Die optimale Dimensionierung ist etwa dann gegeben, wenn der Außendurchmesser das Vierfache des Innendurchmessers beträgt. Eine weitere Verbesserung erhält man dadurch, daß man auf das Stahlrohr eine starkwandige Hülse aufzieht, die dem Rohr eine Vorspannung gibt. Ähnliche Maßnahmen werden bei vielen Druckkammertypen angewendet. Die Wirkung beruht auf folgendem Prinzip: Der Flüssigkeitsdruck im Rohr erzeugt an der Innenwand eine *Zugspannung*. Wenn sie die *Zugfestigkeit des Stahls* überschreitet, reißt die Kammer auf. Durch die aufgeschrumpfte Hülse auf der Außenwand wird an der Innenwand eine Druckspannung erzeugt, die der durch den Innendruck der Flüssigkeit hervorgerufenen Zugspannung entgegenwirkt. Häufig ist die Wand einer Druckkammer aus einem ganzen System von Hülsen bzw. Ringen aufgebaut, deren Wirkungen auf das zentrale Rohr genau abgestimmt sind, so daß bei dem maximalen Innendruck eine gleichmäßige Abnahme der Zugspannung innerhalb der gesamten Kammerwand von innen nach außen erfolgt. Während einlagige Druckkammern nur bis etwa 0,5 GPA benutzt werden, können mit mehrlagigen Zylindern auch 3 GPa realisiert werden. Bei Druckkammern für industrielle Anwendungen mit entsprechend großen Druckräumen wird die Vorspannung auf den zentralen Kammern oft durch Umwickeln mit vorgespannten Drähten aus speziellen Stählen erzeugt.

Eine besondere Bedeutung für die Funktion einer Druckkammer haben die *Dichtungen* der Stempel. Bei ihrer Konstruktion wird häufig vom Prinzip der nichtgestützten Fläche Gebrauch gemacht. Der Stempel besteht dabei aus einem pilzförmigen Oberteil und einem Unterteil mit einer zentralen Bohrung, die den Stiel des Pilzes aufnimmt. Zwischen dem Hut des Pilzes und dem Unterteil befinden sich Ringe aus einem *Dichtungsmaterial*, z. B. aus Gummit, PTFE oder Kupfer. Die Bohrung ist so tief, daß der Pilz nur mit dem Hut unter Zwischenlage der Dichtungen auf dem Unterteil des Stempels aufliegt, der Stiel aber nicht unterstützt wird. In der Druckkammer übt die Druckflüssigkeit eine flächenproportionale Kraft auf die Stirnfläche des Stempels, der Oberseite des Pilzes aus. Mit der gleichgroßen Kraft werden die Dichtungen zusammengedrückt. Da ihre Fläche kleiner als die Stempelstirnfläche ist, entsteht in ihnen ein größerer Druck als in der Druckflüssigkeit. Sie werden dadurch nach außen gegen die Kammerwand gedrückt und dichten diese gegen den Stempel ab. Das Prinzip der nichtgestützten Fläche wird auch bei anderen Konstruktionsformen angewendet. Zur Durchführung von Messungen im Innenraum der Druckkammer wird der nichtbewegte Stempel häufig mit elektrischen Durchführungen versehen.

Es ist möglich, die hydraulische Presse und die eigentliche Druckkammer zu einer konstruktiven Einheit zusammenzufassen. Man kommt auf diese Weise zum *Druckvervielfacher* oder *Multiplikator*.

Bei einer Reihe von Anwendungen ist es erforderlich oder zweckmäßig, eine Druckkammer mit konstantem Volumen zu benutzen. Sie ist in diesem Fall von der eigentlichen Druckerzeugung räumlich getrennt. Die Verbindung von Druckerzeugungssystem und Druckkammer kann über Stahlkapillaren erfolgen, die bis zu Drücken von 1,5 GPa benutzt werden können.

9.3.2.2. Quasihydrostatische Druckerzeugung. Die einfachsten Druckkammern benutzen das *opposed-anvil-Prinzip* (Abb. 9.3.2-3). Das zu drückende Material hat die Form einer flachen Scheibe mit einem Verhältnis des Durchmessers zur Dicke in der Größenordnung von 10 : 1. Diese Scheibe wird zwischen Stempeln aus hochgehärtetem Stahl oder Hartmetall mit hydraulischen Pressen direkt komprimiert. Die Stempel erhalten durch Aufschrumpfen von Ringsystemen eine seitliche Stützung (↑ 9.3.2.1.). Dadurch kann der mit der hohen

Härte gekoppelten Sprödigkeit des Stempelmaterials begegnet werden.
Die eigentliche Probenscheibe wird seitlich von einem Dichtungsring umgeben, der mit der Probe zusammen komprimiert wird. Er verhindert nicht nur die seitliche Extrusion des Probenmaterials (Dichtungsfunktion), sondern wirkt gleichzeitig auch als *Druckübertragungsmittel*. Deshalb soll das Material der Dichtungen bereits bei geringen Drücken in den Bereich des plastischen Fließens kommen, dabei aber eine so hohe innere Reibung haben, daß keine nennenswerte Extrusion auftritt. Außerdem ist eine hohe Kompressibilität erwünscht, damit kein zu großer Bruchteil der Pressenkraft für das Komprimieren der Dichtungen benötigt wird. Diese teilweise widersprüchlichen Anforderungen an ein ideales Dichtungsmaterial sind auch für andere quasihydrostatische Druckkammern zutreffend. Sie werden von wenigen Stoffen annähernd erfüllt, z. B. durch die Minerale Pyrophyllit $Al_2[(OH)_2]$ $[Si_4O_{10}]$, Talk $Mg_3[(OH)_2]$ $[Si_4O_{10}]$, sowie eine Tonvarietät Catlinit.
Kammern nach dem opposed-anvil-System haben den Vorteil der Einfachheit der Konstruktion und Anwendung. Nachteilig ist das geringe Probenvolumen. Sie sind etwa bis in den Bereich von 15 bis 20 GPA einsetzbar, spezielle Ausführungen mit Diamantstempeln für Röntgen- und optische Untersuchungen, sogenannte *Squeezer*, erreichen noch höhere Werte.
Belt- und *Girdle-Kammern* können als logische Weiterentwicklungen der opposed-anvil-Kammern betrachtet werden. Sie besitzen zusätzlich ein Ringsystem, das die seitliche Extrusion der Probe verhindert und eine zusätzliche Stützwirkung auf die Stempel ausübt. Es besteht aus einem zentralen Hartmetallgesenk und einer Reihe von Stahlspannringen, die dem Gesenk eine äußere Stützung geben. Die Dichtungen haben die Form von konischen Hülsen, die die Stempel umgeben. Durch die Verhinderung der seitlichen Extrusion kann die Hülse des Druckraums in der gleichen Größe wie der Durchmesser liegen, das Probenvolumen kann also relativ groß sein (einige Kubikzentimeter). Elektrische Leitungen in den Druckraum können durch die Dichtungen geführt werden. Belt- und Girdle-Kammern können im Druckbereich bis 10 GPa universell eingesetzt werden, z. B. auch bei der industriemäßigen Diamantsynthese (↑ 9.3.9.).

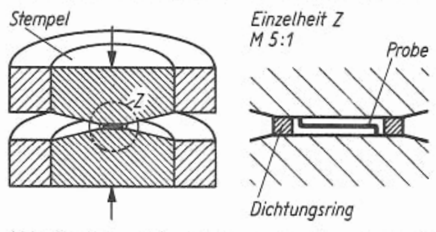


Abb. 9.3.2-3 Prinzip der opposed-anvil-Hochdruckkammer

Tetraederpresse und kubische Presse sind die gebräuchlichsten Formen von Vielstempelkammern (Abb. 9.3.2-4). Bei ihnen hat der Probenraum die Form eines Tetraeders bzw. Würfels. Er wird zur Druckerzeugung dadurch komprimiert, daß auf jede Seite ein gesonderter Stempel einwirkt. Die Anlagen sind relativ aufwendig, weil für den Antrieb jedes Stempels bzw. jedes Stempelpaars eine Presse erforderlich ist. Sie sind aber sehr universell einsetzbar, bis etwa 8 GPa.

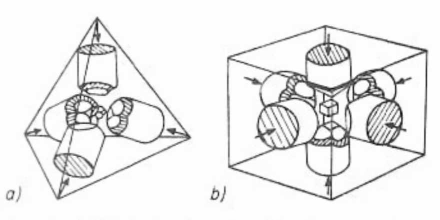


Abb. 9.3.2-4 Prinzip der Vielstempel-Hochdruckkammern; a) Tetraederanordnung, b) kubische Anordnung

Vom Prinzip der Druckerzeugung kann man zu den Vielstempelkammern auch die *Kugelkammern* rechnen. Die Stempel werden hierbei von den Segmenten einer Stahlkugel gebildet, deren Spitzen im Zentrum der Kugel abgeschnitten sind, so daß ein freier Raum entsteht, der den Druckraum darstellt. Es gibt Ausführungen, bei denen die Kugel in sechs Segmente geteilt ist und andere mit einer Unterteilung in acht Segmente bzw. Stempel. Dementsprechend hat der Probenraum die Form eines Würfels oder Oktaeders.

9.3.3. Methoden der Druckmessung

Die Druckmessung unter hydrostatischen Bedingungen führt zu keinen grundlegenden Schwierigkeiten. Bis etwa 0,5 GPa sind direktanzeigende *Bourdonfedermanometer* einsetzbar (Abb. 9.3.3-1). Sie enthalten als druckempfindliches Element eine einseitig geschlossene Röhre mit elliptischem Querschnitt, die zu einem Kreisbogen gespannt ist. Wenn die Röhre mit dem unter Druck stehenden Medium gefüllt wird, streckt sich der Kreisbogen und die entsprechende mechanische Bewegung wird durch eine Übersetzung auf einen Zeiger übertragen. Der Meßfehler liegt günstigstenfalls bei 0,1 % vom Skalenwert.
Im gesamten hydrostatischen Bereich kann

mit *Manganinmanometern* gemessen werden. Der wirksame Drucksensor ist eine nichtinduktive Widerstandsspule aus Manganindraht. Es wird hierbei die Abhängigkeit des Widerstands vom Druck für die Messung ausgenutzt. Die entsprechende Abhängigkeit ist

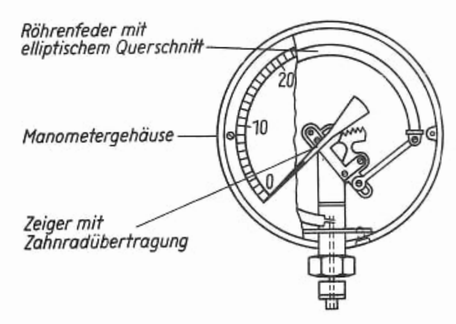


Abb. 9.3.3-1 Bourdonfedermanometer

in guter Näherung linear. Der Druckkoeffizient beträgt etwa $2{,}5 \cdot 10^{-5}\ \mathrm{MPa}^{-1}$. Abweichungen von der Linearität sind bei höheren Genauigkeitsansprüchen oder beim Überstreichen großer Druckintervalle zu berücksichtigen. *Manganin* wurde als Manometerwerkstoff gewählt, weil sein Druckkoeffizient im Vergleich zu anderen Metallen groß ist und der Temperaturkoeffizient des Widerstands niedrig. Es ist erforderlich, das Material vor der Benutzung durch eine thermische Behandlung künstlich zu altern und mehrere Druckzyklen zu durchlaufen, um die erforderliche Reproduzierbarkeit zu erreichen. Da die Spulen auch bei Herstellung aus der gleichen Materialcharge immer etwas unterschiedlich ausfallen, ist bei höheren Ansprüchen eine Eineichung erforderlich. Das Manganin ist gegen mechanische Spannungen sehr empfindlich und muß daher nach Verformungen nochmals gealtert werden. In der praktischen Anwendung werden Spulen mit Widerstandswerten zwischen 10 und 100 Ohm bevorzugt. Für die Widerstandsmessung werden oft Brückenschaltungen benutzt. Bei genauen Messungen macht sich die Temperaturerhöhung beim Komprimieren eines Druckmediums störend bemerkbar.

Wenn die Nichtlinearität bei höheren Drücken durch Korrekturen berücksichtigt wird, können mit Manganinmanometern die Fehler < 0,1% sein. Bourdon- und Manganinmanometer müssen mit Hilfe eines primären Druckmeßgeräts geeicht werden. Dazu werden meist *free-piston-Druckwaagen* benutzt. Das sind im Prinzip kleine Stempel-Zylinder-Kammern, in die die zu messende Druckflüssigkeit eingeleitet wird. Die durch den Innendruck auftretende Druckkraft auf den Stempel wird gemessen, meist durch Kompensation mit Hilfe von Gewichtsstücken. Die Meßgenauigkeit wird entscheidend bestimmt durch die Möglichkeit der genauen Erfassung der elastischen Deformation der Kammer und der Vermeidung der Stempelreibung.

Die *Druckmessung unter quasihydrostatischen Bedingungen* ist sehr problematisch. Die Methoden für hydrostatische Drücke lassen sich bei festen Druckmedien nicht anwenden. Man schließt deshalb von der Größe der druckerzeugenden Kraft auf den Druck. Das Problem bei dieser Verfahrensweise ist, daß ein erheblicher Teil der Kraft für das Komprimieren der Dichtungen aufgewendet wird. Dieser Bruchteil ist unbekannt und im allgemeinen auch druckabhängig, er liegt oft über 50%. Man bewältigt das Problem dadurch, daß man für jede Druckkammer unter Benutzung bekannter *Druckfixpunkte* eine Druck-Kraft-Kalibrierungskurve bestimmt. Die Druckfixpunkte werden durch bestimmte Eichsubstanzen realisiert, die bei festen Druckwerten markante Änderungen ihrer Eigenschaften zeigen. Wegen der einfacheren Meßtechnik sind Substanzen, insbesondere Metalle, für diese Zwecke besonders geeignet, die charakteristische Änderungen (Sprünge oder Extremwerte) der elektrischen Leitwerte zeigen. Einige typische Eichsubstanzen sind mit ihren Eigenschaften in der Tabelle 9.3.3-2 zusammengestellt.

Tab. 9.3.3-2 Wichtige Druckkalibrierungspunkte

Schmelzpunkt von Quecksilber bei 0 °C:			0,76 GPa
Schmelzpunkt von Quecksilber bei 22 °C:			1,18 GPa
I-II-Übergang Bismut	2,55 GPa	I-II-Übergang Blei	13,0 GPa
I-II-Übergang Thallium	3,67 GPa	-- Übergang $Fe_{12}Co$	15,0 GPa
II-III-Übergang Caesium	4,2 GPa	-- Übergang $Fe_{20}Co$	19,0 GPa
I-II-Übergang Barium	5,5 GPa	-- Übergang $Fe_{40}Co$	29,0 GPa
III-V-Übergang Bismut	7,4 GPa	-- Übergang $Fe_{16}V$	38,5 GPa
I-II-Übergang Zinn	9,4 GPa	-- Übergang $Fe_{20}V$	51,0 GPa
α-ε-Übergang Eisen	11,0 GPa		
metallischer Übergang beim Galliumphosphid			22,5

Es gibt nur wenige Fixpunkte über 20 GPa, sie sind außerdem noch relativ unsicher.

Bei röntgenographischen Strukturuntersuchungen unter hohen Drücken wird zur Druckbestimmung der Untersuchungssubstanz Natriumchlorid beigemischt und dann aus der Größe der Elementarzelle des NaCl-Gitters auf den Druck geschlossen. Für die Extrapolation der Druckabhängigkeit der Gitterparameter vom hydrostatischen Bereich zu höheren Drücken ist die Benutzung einer geeigneten Zustandsgleichung erforderlich (↑ 9.3.4.).

Bei Druckmessungen unter quasihydrostatischen Bedingungen sind Fehler von 10% durchaus normal.

Bei Druckkammern, die die Möglichkeit eines optischen Strahlendurchgangs besitzen, kann die Druckabhängigkeit der Lage bestimmter Spektrallinien zur Druckmessung ausgenutzt werden. Man verwendet dabei die sehr intensive R_1-Rubin-Fluoreszenz-Linie, die durch UV-Strahlung anregbar ist und einen Druckkoeffizienten von 0,36 $nmGPa^{-1}$ hat. Dieses Verfahren ist auf Spezialfälle beschränkt und relativ aufwendig. Es ermöglicht jedoch die Extrapolation von niedrigen zu hohen Drücken und damit die Festlegung von Fixpunkten sehr zuverlässig. Dabei wird vorausgesetzt, daß der Zusammenhang zwischen Linienverschiebung und Druck linear ist.

9.3.4. Zustandsgleichungen

Eine Beziehung zwischen dem Druck P, dem Volumen V und der Temperatur T eines Gases, einer Flüssigkeit oder eines Festkörpers nennt man Zustandsgleichung. Dabei ist es üblich, Zustandsgleichungen durch Ableitung eines thermodynamischen Potentials aufzustellen. Kennt man z. B. die *freie Energie* $F = F(T, V)$, so ergibt sich der Druck aus der Beziehung $P = -(\partial F/\partial V)_T$. Auf diese Weise erhält man die sogenannte *thermische Zustandsgleichung*. Es existieren unterschiedliche Möglichkeiten, die freie Energie eines Stoffs zu bestimmen, woraus wiederum verschiedene Zustandsgleichungen resultieren:

Phänomenologische Zustandsgleichungen lassen sich im Rahmen der Kontinuumsmechanik formulieren. Sie sind zur Beschreibung vieler verschiedener Materialien geeignet. Zu dieser Gruppe gehören auch die sogenannten empirischen Zustandsgleichungen von RJABININ und MURNAGHAN.

Mikroskopische Zustandsgleichungen lassen sich z. B. aus der statistischen Theorie des Atoms von THOMAS, FERMI und DIRAC ableiten. Zur Beschreibung verschiedener Festkörpertypen, wie Ionenkristalle, Metalle, Valenzkristalle, verwendet man oft halbmikroskopische Zustandsgleichungen, die empirische Ausdrücke wie n-m-Potentiale für die Paarwechselwirkung zwischen den Gitterbausteinen benutzen. Neben der Auffindung von Zusammenhängen zwischen den einzelnen Arten der Zustandsgleichungen ist die Erforschung solcher Zustandsgleichungen besonders wichtig, die das Verhalten eines Materials in seinen verschiedenen Phasen (↑ 9.3.5.) beschreiben.

9.3.4.1. Zustandsgleichung bei Gasen. Die einfachste Zustandsgleichung ist die eines idealen Gases. Sie lautet $PV = nRT$, dabei ist n die Zahl der Mole und R die Gaskonstante (↑ 6.2.2.1.5.). Als Spezialfälle enthält diese Gleichung das 1. Gay-Lussacsche Gesetz und das Boyle-Mariottesche Gesetz, welche experimentell gefunden wurden. Die Virialform der Zustandsgleichung läßt sich in Form einer Entwicklung nach $1/V$ darstellen. Bei Gasen aus Molekülen ergibt sich für die Zustandssumme ein zusätzlicher Rotations- und Schwingungsanteil. Dadurch resultieren auch entsprechende Korrekturen zur Zustandsgleichung.

Die Van-der-Waals-Gleichung (↑ 6.2.2.2.1.) $(P + a/V^2)(V - b) = nRT$ ist eine Interpolationsformel, die die Eigenschaften eines Gases, einer Flüssigkeit sowie den Phasenübergang zwischen beiden Zuständen qualitativ richtig beschreibt. In Abb. 9.3.4-1 ist die *van-der-Waalssche Zustandsgleichung* in der Nähe der kritischen Temperatur T_k dargestellt.

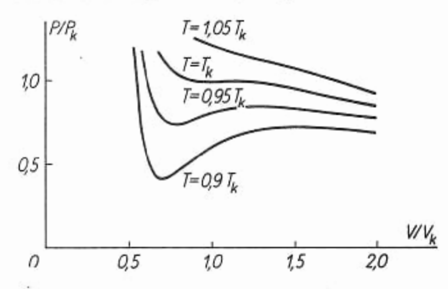


Abb. 9.3.4-1 Van-der-Waals-Gleichung in der Nähe der kritischen Temperatur T_k

9.3.4.2. Zustandsverhalten bei Flüssigkeiten. Die Theorie der Zustandsgleichungen bei Flüssigkeiten ist sehr kompliziert aufgrund der Zwischenstellung zwischen vollkommen ungeordneten Gasen und symmetrisch geordneten Festkörpern. Man wendet deshalb in der Regel Zustandsgleichungen der Gase und Festkörper auch auf Flüssigkeiten an. Befindet sich die Flüssigkeit nahe dem Siedepunkt, kann die Van-der-Waals-Gleichung verwendet werden. Bei Flüssigkeiten in der Nähe des Erstarrungspunkts überwiegen festkörperähnliche Eigenschaften, so daß man bestimmte Zustandsgleichungen für Festkörper anwenden kann.

9.3.4.3. Zustandsgleichungen bei Festkörpern. Zur Beschreibung des *Kompressionsverhaltens* fester Körper verwendet man oft die *Potentialmethode*, die von den verschiedenen Bindungsarten im Festkörper ausgeht. Das Zweiteilchenpotential hat immer einen anziehenden und einen abstoßenden Anteil. Bei *Ionenkristallen* resultiert die Anziehung zwischen den Gitterteilchen aus der Coulombwechselwirkung. Deshalb ist der anziehende Anteil proportional $-1/r$. *Molekülkristalle* werden durch Van-der-Waals-Kräfte zusammengehalten. Hier ist der anziehende Teil proportional $-1/r^6$. *Metalle* werden durch ein Gitter aus positiven Ionen sowie ein Fermigas freier Elektronen charakterisiert. Hier resultiert der anziehende Teil der Wechselwirkung sowohl aus der Coulombwechselwirkung zwischen positiven Ionen und Leitfähigkeitselektronen, als auch aus der Austauschwechselwirkung zwischen den Leitfähigkeitselektronen. Beide liefern den Betrag von $-1/r$. Kennt man das Zweiteilchenpotential eines Festkörpers, läßt sich der Druck $P(V)$ berechnen, d. h., die Zustandsgleichung kann gefunden werden.

Die mit der Potentialmethode erhaltenen Zustandsgleichungen gelten nur für $T = 0$, da die Gitterteilchen als fest angenommen werden. Für $T \neq 0$ tritt neben dem Potentialdruck P_0 noch der sogenannte *Wärmedruck* auf, der von den Gitterschwingungen herrührt. Die Zustandsgleichung wird mit der Form $P = P_0 + \gamma E_{vib}/V$ geschrieben, die sogenannte *Mie-Grüneisen-Gleichung*. Die Größe γ heißt *Grüneisenparameter* und beschreibt die Volumenabhängigkeit der Schwingungsfrequenzen. Wenn man das Verhalten des Grüneisenparameters an einem polymorphen Phasenübergang kennt, läßt sich auch die Mie-Grüneisen-Gleichung zur Beschreibung solcher Phasenübergänge verwenden. Sie hat dann einen ähnlichen Verlauf am Phasenübergang wie die van-der-Waalssche Gleichung.

9.3.4.4. Stoßwellenergebnisse. Komprimiert man ein Material mit Hilfe von Stoßwellen, lassen sich die Ergebnisse in Form der *Hugoniotschen Zustandsgleichung* $P'/P - V/V' = [(\varkappa - 1)/2]\,[(1 + P'/P)\,(V/V' + 1)]$ darstellen. P' und V' bedeuten Druck und Volumen nach dem Stoß, $\varkappa$ ist der Index der Adiabaten $P'/P = (V/V')^{\varkappa}$. Bei Kompressionen bis etwa $V'/V = 1/2$ weicht die Hugoniotkurve nur sehr wenig von der Adiabaten ab.

Durch Schockwellen lassen sich sehr hohe Drücke erzielen. Deshalb kommt es bei den meisten der untersuchten Materialien zu Phasenübergängen in dichtere Phasen. Solche Phasenübergänge führen zu Unstetigkeiten in der Hugoniotkurve.

Für eine eindeutige Aufklärung des Phasenverhaltens eines Stoffs ist der Vergleich der Hugoniotkurve mit *isothermen* und *adiabatischen Kurven* notwendig. Letztere werden z. B. aus Hochdruckröntgenexperimenten und Ultraschalluntersuchungen gewonnen. Bei genauer Kenntnis des Grüneisenparameters ist eine Reduktion auf adiabatische und isotherme Bedingungen mit Hilfe der Mie-Grüneisen-Gleichung möglich.

Abb. 9.3.4-2 zeigt die Hugoniotkurve für Forsterit (Mg_2SiO_4). Der Übergang in die dichtere Spinellphase ist deutlich sichtbar.

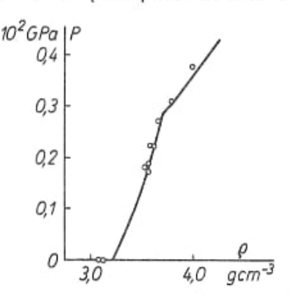


Abb. 9.3.4-2 Hugoniotkurve für Forsterit; P Druck, ϱ Dichte

9.3.5. Phasenumwandlungen bei hohen Drücken

Jede thermodynamische stabile Phase ist durch ein Stabilitätsfeld charakterisiert, das durch die Größen Druck p, Temperatur T und Zusammensetzung x bestimmt ist. Im p, T, x-Diagramm können Übergänge von einer Phase in die andere auch eindimensional in Richtung des Drucks erfolgen. Derartige rein druckinduzierte Phasenumwandlungen können Übergänge der Art flüssig $\leftrightarrow$ fest oder fest $\leftrightarrow$ fest sein. So ist zu beobachten, daß im Druckbereich bis 10 GPa alle unter Normalbedingungen bekannten Flüssigkeiten in eine feste Phase übergehen. Übergänge fest $\leftrightarrow$ fest unter dem Einfluß hoher Drücke sind meist *polymorphe Phasenumwandlungen* und führen zu Hochdruckmodifikationen bekannter kristalliner Phasen. Diese Umwandlungen können reversibel oder irreversibel sein. Die bekannteste irreversible polymorphe Phasenumwandlung ist die Drucktransformation von Graphit in Diamant, der dabei metastabil erhalten bleibt. Entsprechend ihrem Mechanismus unterscheidet man außerdem nach BUERGER *rekonstruktive* und *displazive Umwandlungen*. Erstere erfordern die Überwindung einer Energiebarriere ΔE, da primäre Bindungen gelöst werden müssen und die Struktur völlig neu aufgebaut wird.

Die displazive Umwandlung beinhaltet eine Verdrehung, Verschiebung oder Umordnung von Atomgruppen oder einer größern Zahl benachbarter Atome, wodurch die neue Struktur entsteht.

Phasenumwandlungen laufen unter Änderung der *freien Enthalpie* ΔG ab. Bei druckinduzierten Umwandlungen zeigt sich dabei eine negative Volumenänderung $-\Delta V$. Hochdruckmodifikationen fester Phasen werden in der Regel dichter und härter sein sowie bessere Leitfähigkeit zeigen.

9.3.5.1. Strukturübergänge in kristallinen Festkörpern. Der experimentell beherrschbare Druck reicht bereits bis in den 100-GPa-Bereich (↑ 9.3.1.), und in diesen weiten Druckgrenzen sind bei fast allen Stoffen bereits Strukturübergänge unter Druck nachgewiesen worden.

Dabei sind einige Regeln entdeckt worden, die von NEUHAUS formuliert wurden:

Druckkoordinationsregel: Sie besagt, daß mit steigendem Druck die Koordinationszahl (KZ) steigt, d. h., bei Strukturübergängen unter Druck werden dichter gepackte Strukturen mit höherer KZ realisiert (Tab. 9.3.5-1).

Druckhomologieregel: Sie besagt, daß niedere Homologe in homologen Reihen des PSE unter Druck die Strukturen ihrer höheren Homologen erreichen. Gute Beispiele dafür liefern die Alkalihalogenide, die IV. Hauptgruppe des PSE Sn → Ge → Si → C sowie zahlreiche Silicate mit homologen Kationen.

Damit verbunden ist die Änderung des Valenzzustands, der die elektrische Leitfähigkeit bestimmt. Man kann heute schon voraussagen, daß alle Substanzen durch Druckeinwirkung Strukturübergänge durchlaufen, die sie letztlich in den metallischen Zustand überführen. Die druckbedingten strukturellen Phasenübergänge spielen für die Modellvorstellungen zum Aufbau des tiefen Erdinneren eine entscheidende Rolle (↑ Abb. 9.3.4-2).

Technische Bedeutung haben druckinduzierte strukturelle Phasenübergänge auch für die Werkstofforschung (z. B. Synthese von Diamantstaub und kubischem Bornitrid; ↑ 9.3.9.).

9.3.5.2. Schmelzen. Das Schmelzen ist ein spezieller Phasenübergang eines Materials von der festen in die flüssige Phase. Wirkt auf einen Festkörper ein Druck, erhöht sich im allgemeinen die *Schmelztemperatur* T_s. Der Grund liegt in der etwas geringeren Dichte der Flüssigkeit. Die Atome sind im Festkörper relativ dicht gepackt; sie vollführen thermische Schwingungen um ihre Ruhelagen. Erhöht man die Temperatur, vergrößern sich die Schwingungsamplituden, so daß sie, wenn der Schmelzpunkt erreicht wird, ihre Positionen verlassen und gegeneinander verschoben werden, man spricht von Dislokationen. Dadurch besitzt die flüssige Phase eine etwas geringere Dichte als die entsprechende feste. Man kann also Flüssigkeiten durch isotherme Kompression einfrieren.

Eine Ausnahme bilden Materialien, die im

Tab. 9.3.5-1 Beispiele zur Druckkoordinationsregel

	Normaldruckmodifikation				Umwandlung		Hochdruckmodifikation		
Formel	Name	Symmetrieklasse	Gittertyp		10^8 Pa	°C	Kristallklasse	Gittertyp (Name)	KZ
C	Graphit	$6/mm$	Graphit	3 + 1	≧60	≧1250	$m3m$	Diamant	4
Sn	Zinn, weiß	$4/mm$	weißes Zinn	6	>115	20	$m3m$	kub. raumz.	8
BN	Bornitrid	$6/mm$	Graphit	3 + 1	≧60	1300	kubisch	(Borazon)	4
SiO_2	Quarz	32	Quarz	4	20	500	$2/m$	(Coesit)	4
SiO_2	Coesit	$2/m$		4	125	≧1200	$4/m$	Rutil (Stishovit)	6
NaCl	Steinsalz, Natriumchlorid	$m3m$	NaCl	6	≧100	20	$m3m$	CsCl	8
AgCl	Chlorargyrit, Silberchlorid	$m3m$	NaCl	6	85	20	$m3m$	CsCl oder Hg_2Cl_2	8
AlAl$[SiO_5]$	Sillimanit	mmm		Al: 6 + 4	17	1200	1	$Al_2[O/SiO_4]$	Al: 6 + 6
$CdTiO_3$	Cadmiumtitanat	3	Ilmenit	6(Cd)	12	600	$m3m$	Perowskit	12 (Cd)
$CaCO_3$	Calcit	$3m$	Kalkspat	6	6	20	mmm	Aragonit	9

festen Zustand eine sehr „lockere“ Struktur haben, oder bei denen das Schmelzen mit einem anderen Phasenübergang verbunden ist, so das letztlich die flüssige Phase eine größere Dichte besitzt als die feste. Zu diesen Ausnahmen gehört das Wasser.
Eine Methode, die Abhängigkeit der Schmelztemperatur T_s vom Druck P zu berechnen, basiert auf der *Lindemannschen Regel.* Diese besagt, daß ein Kristallgitter schmilzt, wenn die Amplituden der Gitterschwingungen der Atome um ihre Ruhelagen einen kritischen Bruchteil des Abstands zwischen den nächsten Nachbarn erreicht. Wird infolge von Kompression das Volumen verkleinert, reduzieren sich die Abstände zwischen den Atomen des Gitters. Die Regel besagt außerdem noch, daß sich dann auch die kritischen Schwingungsamplituden in dem gleichen Maße verkleinern. In dieser Weise kann man versuchen $T_c(P)$ abzuschätzen. Da in die Berechnung jedoch mehrere schwer bestimmbare Größen eingehen (z. B. die interatomaren Potentiale) ist das Verfahren noch recht unsicher.

9.3.6. Elektrische Eigenschaften

9.3.6.1. Elektrolyte. Der Druckeffekt auf den *elektrischen Widerstand* von Elektrolyten ist sehr kompliziert. Bei normalen Temperaturen und in verdünnten Lösungen starker Elektrolyte bewirkt er eine Abnahme des elektrischen Widerstands. Mit weiterer Steigerung des Drucks verflacht dieser Effekt. Der Druckeffekt ist stark konzentrationsabhängig. Weiterhin spielt die Art des Elektrolyten eine große Rolle. So kann der elektrische Widerstand zunächst ansteigen, um dann bei Steigerung des Drucks wieder abzunehmen. In Richtung sehr niedriger Elektrolytkonzentrationen (im Wasser) verschwindet der Druckeffekt, weil dann die Eigenleitfähigkeit des Wassers wirksam wird. Bei schwachen Elektrolyten vergrößert sich der anfängliche Druckeffekt sehr stark mit der Konzentration. Bei Konzentrationen unter 0,001 n ist er minimal größer, als bei darüber liegenden Konzentrationen.
Für dieses Verhalten gibt es insgesamt vier wichtige Effekte, die einen Beitrag liefern: Der reine Volumeneffekt, der mehr Ionen in eine Volumeneinheit zwingt, die Beweglichkeit der Ionen, die in erster Näherung parallel zur Viskosität verläuft, die Druckauswirkung auf den Grad der Dissoziation sowie der Druckeffekt auf die Dissoziation bezüglich der Eigenleitfähigkeit des Wassers.
Die beiden letzten Effekte können bei starken Elektrolyten vernachlässigt werden. Der erste Effekt kann aus den Kompressibilitäten bestimmt werden, bringt bisher aber sehr widersprüchliche Ergebnisse. Die Beweglichkeit der Ionen kann ebenfalls nicht exakt erfaßt werden. In vielen Fällen sinkt sie zunächst, steigt bei höheren Drücken aber wieder an. Diese Änderungen überdecken Größenordnungen und hängen stark von der Art des Elektrolyten ab.
9.3.6.2. Halbleiter. Halbleitende Substanzen gehen bei entsprechenden Drücken in den *metallischen Zustand* über. Elemente der Gruppe IV und binäre gleichatomige Verbindungen zwischen den Gruppen III und V transformieren bei hohen Drücken von der Diamant- oder Zinkblendestruktur in den metallischen Zustand mit der Struktur des weißen Zinns. Der metallische Zustand wird auch bei Verbindungen aus den Elementen der Gruppen II und VI erreicht. Aus theoretischen Überlegungen (z. B. Theorie der kovalenten Radien) wurden z. B. die Übergangsdrücke für Ge, Si, GaAs, AlSb, GaSb in guter Übereinstimmung mit dem Experiment bestimmt (Tab. 9.3.6-1).

Tab. 9.3.6-1 Übergangsdrücke in verschiedenen Verbindungen

Verbindung	Übergangsdruck in GPa
GaP	50
BP	40
BN	226
AlN	90
GaN	87
SiC	64

Der *druckinduzierte Übergang* zur metallischen Leitfähigkeit geht meist sehr rapide vor sich. Im Falle des GaP oder CdSe steigt die Leitfähigkeit innerhalb einer Minute (im Experiment) um sechs bis sieben Größenordnungen, wenn der Übergangsdruck erreicht ist.
Bei BN, AlN und GaN verringert sich der elektrische Widerstand monoton mit der Steigerung des Drucks.
9.3.6.3. Metalle. Elektrische Widerstandsmessungen in Abhängigkeit vom Druck wurden zuerst an relativ inkompressiblen Metallen durchgeführt, z. B. Fe, Ag, Cu und Pt. Allgemein sinkt bei diesen Metallen der elektrische Widerstand mit steigendem Druck. Bis etwa 1 GPa entspricht das jedoch nur 1 bis 2% des Ausgangswerts. Bei kubisch kristallisierenden Metallen ist die Leitfähigkeitsänderung richtungsunabhängig. Treten druckbedingte Phasenwechsel auf, kann es auch zu sprunghaften positiven Druckkoeffizienten kommen, z. B. bei Bismut. Bei Legierungen ist das Leitfähigkeitsverhalten wesentlich komplizierter.

9.3.6.4. Supraleiter. Ein wichtiges Ziel der Hochdruckforschung besteht im Auffinden neuer Hochdruckmodifikationen mit supraleitenden Eigenschaften. Weniger als 30 Elemente zeigen bei Normaldruck supraleitende Eigenschaften. Unter Hochdruckbedingungen bis 20 GPa kommen jedoch noch 14 weitere Elemente hinzu, die bei Normalbedingungen keine Supraleiter darstellen. Dazu gehören Silicium, Germanium, Selen, Tellur, Phosphor, Arsen, Cerium und Caesium.
Die tetraedrisch gebundenen kristallinen Halbleiter Silicium und Germanium wandeln sich bei Drücken von etwa 13 GPa bzw. 10 GPa in eine Modifikation mit der Kristallstruktur des α-Zinns um. Diese Modifikationen haben metallische Eigenschaften und können supraleitend werden.
Weiterhin können Hochdruckmodifikationen von Elementen, die bereits bei Normaldruck supraleitend sind, höhere Sprungtemperaturen als die bei Normaldruck stabile Modifikation besitzen. Beispiele dafür sind Zinn und Gallium.
Theoretische Untersuchungen lassen erwarten, daß metallischer Wasserstoff supraleitende Eigenschaften besitzt. Es wurden *Sprungtemperaturen* in der Größenordnung 100 bis 300 K abgeschätzt. Sollte metallischer Wasserstoff auch bei Normaldruck im metastabilen Zustand mit diesen Eigenschaften existieren können, würde das zu einer Revolution in der gesamten Technik führen.

9.3.7. Spektroskopische Untersuchungen

Die experimentelle Technik erlaubt heute die Aufnahme der verschiedensten Spektren in Abhängigkeit vom Druck wie Absorptions- und Emissionsspektren, IR-Spektren, Ramanspektren.
Der Druckeffekt, der sich dabei in Flüssigkeiten oder in Festkörpern zeigt, ist abhängig von der erreichten Dichteänderung. Die beobachteten Linienverschiebungen im Spektrum liefern dabei vielfältige Informationen:
Die Verschiebung der Absorptionsgrenze unter Druck erlaubt die Berechnung des Bandabstands zwischen *Leitungs- und Valenzband* im Festkörper.
Spektren von Übergangsmetallkomplexen unter Druck liefern Ligandenfelddaten und zeigen Änderungen im Bindungsverhalten. Hieraus lassen sich Energien zwischen Elektronenniveaus berechnen und der Anteil kovalenter Bindung bestimmen, der mit dem Druck steigt.
Das Auftreten von druckinduzierten Linienverbreiterungen und das Erscheinen neuer Linien im Spektrum erlaubt den Nachweis von Ladungsänderungen in Kationenkomplexen unter Druck. Emissionsspektren unter Druck lieferten z. B. Zerfallsraten von Phosphorverbindungen in Abhängigkeit vom Druck (nach DRICKAMER).
UV- und IR-Spektren, aufgenommen in Abhängigkeit vom Druck, finden breite Anwendung für die Defektanalyse in Kristallen zur Untersuchung von Störstellen, Farbzentren u. a. Weitere Anwendungen spektroskopischer Untersuchungen unter Druck bestehen z. B. im Nachweis elektronischer Übergänge sowie in der Berechnung von Phononenspektren.

9.3.8. Magnetische Eigenschaften

Die interatomaren Kräfte bewirken die Ausrichtung der Atommomente und damit die Einstellung des magnetischen Typs der Substanz in Form von *Ferromagnetismus, Antiferromagnetismus oder Ferrimagnetismus.*
Ein Maß für diese interatomaren Wechselwirkungen sind die *Curie-* bzw. *Néeltemperaturen* der Stoffe. Die Druckabhängigkeit dieser Grenztemperaturen der magnetischen Zustände bilden den Schwerpunkt magnetischer Messungen unter Druck.
Experimentell wird dabei die *magnetische Suszeptibilität* oder die Sättigungsmagnetisierung im magnetfreien Druckraum, bevorzugt unter hydrostatischen Druckbedingungen, bestimmt.
Ferromagnetismus hängt direkt von den interatomaren Abständen ab und ist deshalb besonders druckempfindlich. Da der Druck das Kristallgitter verändert, lassen sich Aussagen zum magnetischen Charakter einer Substanz über die Druckabhängigkeit der magnetokristallinen Anisotropie gewinnen, der Größe, die die Kopplung zwischen den magnetischen Momenten der Atome und den Richtungen im Kristallgitter beschreibt.

9.3.9. Hochdrucksynthesen

Bis Anfang der 50er Jahre wurde der Parameter Druck fast völlig bei der Synthese von Materie im festen Zustand ignoriert, d. h. stillschweigend als konstant angenommen. Das änderte sich schlagartig als 1953 die erste *Diamantsynthese* gelang und damit auch neue Impulse für die Beherrschung derart extremer thermodynamischer Bedingungen gegeben werden konnten. Inzwischen sind hohe Drücke für die Synthese verschiedenster Materialien sowie die Untersuchung deren Phasenverhaltens selbstverständlich. Es sollen nachfolgend nur die wichtigsten Synthesen – auch unter technischen Aspekten – kurz beschrieben werden.

9.3.9.1. Diamant. Die erste seriöse Arbeit über eine angeblich erfolgreiche *Diamantsynthese* erschien 1880 von HANNAY, der Hydrocarbonate mit Lithium und Öl in Eisenröhren erhitzte, die dann durch den inneren Druck auseinanderbarsten. Die Produkte hatten eine Dichte von 3,5 gcm^{-3} und enthielten 97% Kohlenstoff. Die ersten Druckversuche führte

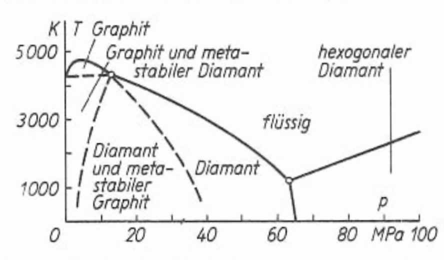


Abb. 9.3.9-1 Zustandsdiagramm des Kohlenstoffs

1894 MOISSON in Autoklaven durch. Er löste Zuckerkohle in einer Eisenschmelze und schreckte beides in Wasser ab. Durch den inneren Druck beim Abkühlen sollte sich die Umwandlung vollziehen. Solche Erfolgsmeldungen wurden anschließend mit der Einführung der Röntgenstrukturuntersuchungen seltener. Es verstärkte sich die theoretische Diskussion über die thermodynamischen Bedingungen und Möglichkeiten der Synthese. Dabei konnten 1955 BERGMANN und SIMONS die Transformationskurve Graphit – Diamant berechnen (Abb. 9.3.9-1). Außerdem konnte aus reaktionskinetischen Betrachtungen der Beweis angetreten werden, daß eine metastabile Diamantbildung im Sinne der Ostwaldschen Stufenregel für das reine Einstoffsystem aussichtslos ist. Dem schienen die Versuche von BRIDGMAN zu widersprechen, der Diamant zwischen 1,5 und 3,0 GPa bei 3000 °C in Graphit umwandelte. Er beobachtete eine starke Abnahme der Umwandlungsrate mit dem Druck. Bei 3 GPa und 3000 °C war keine Umwandlung mehr festzustellen. Auf dieser theoretischen Grundlage gelang 1953 fast gleichzeitig zwei Forschergruppen in den USA und in Schweden die erste Synthese von Diamant. In beiden Fällen wurden die Synthesen – ausgehend von bestimmten Metall-Kohlenstoff-Systemen – bei Drücken zwischen 6 und 9 GPa und Temperaturen zwischen 1600 und 2000 °C erzielt. Bis zur Gegenwart konnten diese Synthesebedingungen weiter zu geringeren Temperaturen und Drücken verbessert werden. In Abhängigkeit von der gewünschten Korngröße und den eingesetzten Katalysatoren werden bei der *industriellen Diamantsynthese* gegenwärtig etwa 5,5 GPa bei 1400 bis 1500 °C angewendet. Der Synthese liegt dabei die von HALL erstmals eingesetzte Hochdruckkammer vom Typ „Belt" (↑ 9.3.2.2.) zugrunde. Zu Forschungszwecken gelangen Synthesen auch mit Hilfe anderer Druckkammertypen sowie mit Schockwellen.

Gegenwärtig werden in über 20 Ländern Diamanten synthetisch hergestellt. Der steigende Weltverbrauch bestätigt die steigende Nachfrage und Erweiterung der *Einsatzgebiete.* Im Jahre 1970 wurden mehr als 70 Mill. Karat produziert. Dabei muß berücksichtigt werden, daß bei der großtechnischen Diamantproduktion nahezu ausschließlich Körnungen von maximal 0,5 mm gewonnen werden. Diese Produkte kommen zum Schleifen und Polieren oder in Metall- und Kunstharzbindungen in Form von Schneid- und Trennscheiben zum Einsatz. Zur Bestückung von Einzelkornwerkzeugen, wie sie zum Bohren, Fräsen oder Sägen benutzt werden, setzt man bisher nur große Naturdiamanten ein. Die Synthese großer preiswerter Diamanten wird bisher serienmäßig nicht beherrscht. Eine Alternative zu großen Diamanteinkristallen bilden für viele technische Anwendungen harte und verschleißfeste *polykristalline Diamantkörper*, die in geringen Mengen auch in der Natur als *Carbonados* vorkommen.

Bei Betrachtung der Kohlenstoffgitter scheint die Umwandlung Graphit – Diamant völlig problemlos zu sein (Abb. 9.3.9-2), da rein geometrisch eine Kompression der c-Achse des Graphitgitters zur kubischen Diamantstruktur führt. Daraus resultierte dann auch die falsche Annahme, daß bei genügend hohem Druck allein durch den Kompressionseffekt eine Umwandlung möglich sei. Versuche von DRICKHAMER zeigten, daß Graphit bei 18 GPa noch 84% und bei 30 GPa immer noch 82% seines Ausgangsvolumens einnimmt. Die Diamantstruktur besitzt demgegenüber nur 64% des Graphitvolumens. Es ließ sich schließlich ableiten, daß sich die Umwandlung über eine vollständige Rekonstruktion des Gitters und das Lösen von Primärbindungen vollzieht. Die Praxis hat inzwischen beweisen können, daß bei diesen Vorgängen die Mitwirkung eines Katalysators beträchtliche Vorteile bietet. Obwohl der eigentliche Katalyseprozeß gegen-

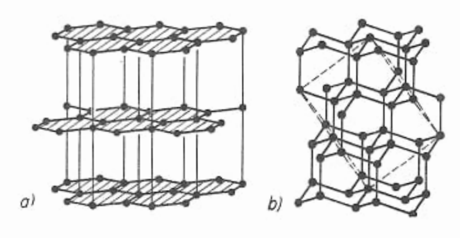


Abb. 9.3.9-2 Kristallgitter a) von Graphit, b) von Diamant

wärtig noch weitgehend unklar ist, weiß man, daß bei der Synthese die Löslichkeit von Graphit in metallischen oder halbmetallischen Schmelzen ausgenutzt wird. Man unterscheidet z. B. graphiterzeugende und diamanterzeugende Lösungen. Zur zweiten Gruppe gehören Metalle wie Mangan, Eisen, Cobalt, Nickel, Platin und Kupfer, die auch bei der industriemäßigen Synthese eingesetzt werden.

Die systematische Untersuchung des *Phasenverhaltens von Kohlenstoff* unter extremen Druck- und Temperaturbedingungen führte bei 16 GPa und etwa 1000 °C zum Nachweis der *hexagonalen Diamantmodifikation.* Eine Phasengrenze zwischen der kubischen und der hexagonalen Modifikation konnte nocht nicht gefunden werden. Interessant ist, daß hexagonaler Diamant (Longsdalit) in Meteoriten nachgewiesen werden konnte. Durch Schockwellen fand man auch eine *metallische Modifikation* des Kohlenstoffs.

Den ständig steigenden Einsatz auf vielen Gebieten der Technik – die Verwendung als Schmuckstein ist demgegenüber verschwindend gering – verdankt der Diamant einer Fülle einzigartiger und äußerst nützlicher *physikalischer Eigenschaften.* Im Vordergrund steht zweifellos die außerordentliche *Härte* und große chemische Beständigkeit. Ursache dafür ist das kubische Diamantgitter mit der in allen Richtungen gleichmäßigen homöopolaren Bindungsenergie von etwa 60 kcalMol^{-1} je C—C-Bindung. Nach der Knoopskala hat Diamant eine Härte von maximal 8500 in Abhängigkeit von der kristallographischen Orientierung. Die Werte für Hartstoffe wie Wolframkarbid liegen bei maximal 3000. Diamant besitzt eine hohe Verschleißfestigkeit (E-Modul zwischen 720000 und 1000000 Nmm^{-2}) und ist am wenigsten kompressibel. Auch seine *thermischen Eigenschaften* sind sehr ungewöhnlich: Die Wärmeleitfähigkeit entspricht der von Kupfer größenordnungsmäßig. Dotierte Diamantkristalle zeigen *Halbleitereigenschaften* und können zum Bau von Thermistoren und Transistoren verwendet werden. Mit solchen Bauelementen sind Arbeitstemperaturen bis 500 °C möglich.

Hinsichtlich seiner Lichtabsorption im infraroten und ultravioletten Spektralbereich unterscheidet man verschiedene Typen von Diamant. Von seinen hervorrage den *optischen Eigenschaften* sind die hohe Lichtbrechung und die starke Dispersion besonders bekannt, die den geschliffenen Diamant als Brillant zu einem begehrten Schmuck werden ließen.

9.3.9.2. Bornitrid. Bestimmte Homologien in der Geometrie und anderer Eigenschaften zwischen Graphit und hexagonalem Bornitrid ließen bereits in den Anfängen der Diamantsynthese eine kubische Bornitridmodifikation mit ähnlichen hervorragenden Eigenschaften vermuten. Anfang der 60er Jahre gelang dann auch die Synthese von kubischem Bornitrid. Die hochdrucktechnologischen Voraussetzungen sind die gleichen wie bei der Diamantsynthese, die notwendigen Drücke und Temperaturen liegen mit etwa 6 GPa und 1600 °C jedoch höher als beim Diamant. Hinsichtlich seiner Verwendbarkeit bietet Bornitrid den Vorteil, daß es thermisch stabiler ist, was sich besonders günstig auf die Standzeiten der entsprechenden Werkzeuge auswirkt. Sowohl die Härte als auch die Möglichkeit polykristalline Körper herzustellen, bieten für Bornitrid die gleiche Einsatzbreite in der Technik.

9.3.9.3. Quarz. Die technische Beherrschung hoher Drücke hat die systematische Untersuchung des Phasenverhaltens der verschiedensten Substanzen stimuliert und damit die Kenntnis des *Druck-Temperatur-Felds* stark erweitert. Wie Abb. 9.3.9-3 zeigt, konnte z. B. beim Quarz eine Reihe neuer Phasen synthetisiert werden, die teilweise inzwischen auch in der Natur nachgewiesen werden konnten. Dies hat sowohl für die geowissenschaftliche Forschung und Anwendung wie auch für viele technische Anwendungen entscheidende Konsequenzen.

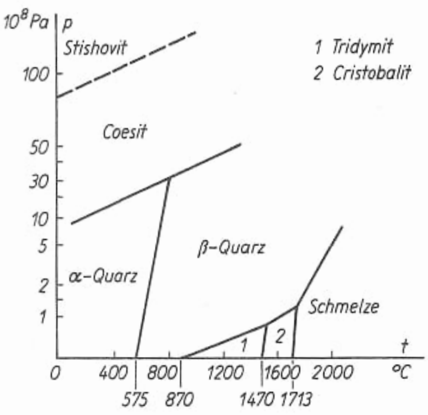


Abb. 9.3.9-3 Phasendiagramm des SiO_2; *p* Druck, *t* Temperatur

9.3.9.4. Hydrothermalsynthesen. Für eine Vielzahl von vorwiegend anorganischen Substanzen stellt die Hydrothermalsynthese die günstigste bzw. einzige Methode zur Herstellung ausreichend großer Kristalle dar. Eine akzeptable Ausbeute ist jedoch bei den meisten nur unter Hochdruckbedingungen möglich. Für diesen Zweck werden *Autoklaven* eingesetzt, die es – je nach eingesetzter Substanz – gestatten, bei Drücken von einigen 100 MPa und mehreren 100 °C zu arbeiten. Zur Zeit hat die Herstellung künstlicher großer Quarzkristalle

größte Bedeutung. Das gilt auch für die Herstellung synthetischer Edelsteine. Es gelingt z. B. Smaragdkristalle bei 1 GPa und 1800 °C zu synthetisieren. Dabei werden die Ausgangskomponenten SiO_2, Al_2O_3 und BeO unter Zusatz von Chromoxid in einer Beltapparatur unter den genannten Bedingungen zur Reaktion gebracht.

Die Autoklaventechnik wird schließlich auch verstärkt für die Untersuchung *technischer Gläser* eingesetzt. Dabei ist es durch Druck prinzipiell möglich, spezifische Eigenschaften wie Brechwert oder Dispersionsverhalten gezielt zu beeinflussen. Schwierigkeiten bereitet demgegenüber noch die gesteuerte Kristallisation technischer Gläser.

9.3.9.5. Chemische Reaktionen unter hohem Druck. In der chemischen Industrie werden bei vielen technisch wichtigen Verfahren hohe Drücke angewendet. Sie dienen dabei grundsätzlich zur Beeinflussung der *Reaktionsgeschwindigkeit* bzw. der Lage des Gleichgewichts mit dem Ziel der Erhöhung der Ausbeute. Besonders typisch dafür sind die Ammoniaksynthese und die Herstellung von Hochdruckpolyethylen. Es muß jedoch für jedes Produkt speziell untersucht werden, ob der Einsatz hohen Drucks prinzipiell und auch ökonomisch zweckmäßig ist. Dabei ist außerdem zu berücksichtigen, daß für die großtechnische Produktion entsprechend große Abmessungen der Reaktionsgefäße erforderlich sind. Dann nämlich wirken so hohe Kräfte auf die Wandungen, daß schließlich nur relativ geringe Drücke eingesetzt werden können. In der Regel liegen die Drücke selten über 100 MPa.

9.3.10. Isostatisches Pressen

Unter isostatischem Pressen versteht man die *Kompaktation* einer Pulvermenge in einer Flüssigkeit oder einem Gas durch Druckerhöhung, wodurch eine feste, isotrope Masse entsteht. Dabei besteht die Möglichkeit einer gewünschten Formgebung des Preßprodukts. Man unterscheidet zwischen dem isostatischen *Kaltpressen* und dem isostatischen *Heißpressen*, dem sogenannten Drucksintern.

9.3.10.1. Isostatisches Kaltpressen. Das Verfahren hat gegenüber dem einfachen konventionellen mechanischen Pressen den Vorteil, daß die Preßlinge eine höhere Homogenität besitzen. Außerdem können in einfacherer Weise recht komplizierte Teile hergestellt werden, wie Hohlkörper mit Längsnuten, Gewinde, Turbinenschaufeln. Die Anwendung konzentriert sich auf die Pulvermetallurgie (z. B. Hartstoffe), die keramische Industrie (z. B. Hochspannungskeramik, Porzellan), die Formung von Sprengstoffen und Kunststoffen sowie die Herstellung von Kohleformkörpern (z. B. Elektroden). Durch das isostatische Pressen können bei Oxidkeramik die Bindemittel reduziert und dadurch bessere elektrische Eigenschaften erzielt werden. Das Verfahren wird bei Normaltemperatur bis zu Drücken von etwa 0,5 GPa eingesetzt.

9.3.10.2. Isostatisches Heißpressen. Bei diesem Verfahren wird gegenüber dem Kaltpressen mit einem Inertgas gearbeitet. Die Temperaturen liegen bei maximal 1800 °C, der Druck beträgt maximal 0,3 GPa. Mit diesem weitaus aufwendigeren Verfahren erreicht man für die eingesetzten Materialien eine noch höhere Dichte, die bei bestimmten Teilen unbedingt erforderlich ist (z. B. Laufbuchsen, Walzen).

9.3.11. Hydrostatisches Strangpressen

Dieses auch unter *Extrusion* bekannte Verfahren der *Werkstoffverformung* gehört gleichfalls zu den ökonomisch intensivsten Anwendungen hoher Drücke in der Technik. Dabei wird das zu bearbeitende Material – meist in Stangenform – in einer Hochdruckkammer (Tafel 48) über die Elastizitätsgrenze gebracht und im plastischen Zustand durch eine Düse in die gewünschte Form gepreßt. Abb. 9.3.11-1 zeigt das Grundprinzip des Verfahrens im Unterschied zu dem des konventionellen Strangpressens, wo die Preßkraft von einem Stempel direkt auf den Rohling übertragen wird.

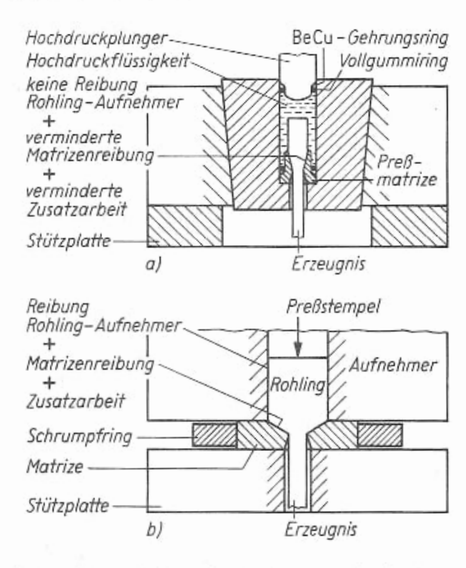


Abb. 9.3.11-1 Grundprinzipien a) des hydrostatischen Strangpressens, b) des konventionellen Fließpressens

Die angewendeten Drücke richten sich nach den Eigenschaften des eingesetzten Metalls und liegen maximal bei 1,5 GPa. Das Verfahren wird auch bei Stahl und legierten Stählen angewendet. Abb. 9.3.11-2 vermittelt einen Eindruck der Leistungsfähigkeit dieses Verfahrens. Besonders bietet sich das Verfahren zum Ziehen feinster Drähte bis 0,015 mm an.

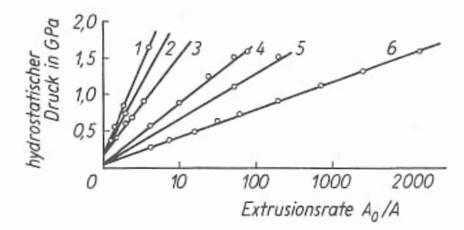


Abb. 9.3.11-2 Extrusionsraten verschiedener Materialien; *1* Schnelldrehstahl, *2* Stahl 0,35 C, *3* Stahl 0,15 C, *4* Kupfer 99,9%, *5* AlCuMg-Legierung, *6* Aluminium 99,5%

9.3.12. Explosionsumformung

Es existiert gegenwärtig bereits eine Vielzahl spezieller Verfahren zur Beeinflussung metallischer Werkstoffe mit Explosivstoffen. Dazu gehören das Explosivschweißen, Explosivhärten, Stoffumwandlungsverfahren mit entsprechenden Hohlladungen und Sprengpressen von metallischen und keramischen Pulvern. Unter diesen Verfahren nimmt das *Explosivschweißen* oder Explosivplattieren zweifellos den größten Raum ein. Damit lassen sich metallische Werkstoffe, auch wenn sie als unverschweißbar gelten, innig miteinander verbinden. Der Mechanismus besteht im Erzeugen einer Druckwelle über der Verbindungsfläche durch die Detonation des Sprengstoffs. Die hohe Umformgeschwindigkeit von 5000 bis 12000 ms^{-1} und der Druck von 20 bis 40 GPa bei Detonation des Sprengstoffs erzeugen in der Verbindungsfläche eine so hohe Energie, daß sich der Werkstoff im submikroskopischen Bereich wie ein Plasma verhält. Anwendung findet dieses Verfahren z. B. beim Aufbringen von dünnen Korrosionsschutzschichten.

9.3.13. Anwendungen hoher Drücke in den Geowissenschaften

Die Erforschung des strukturell-stofflichen Aufbaus der Erde und der Planeten erfordert die Einbeziehung des Drucks in der mannigfaltigsten Weise.

Bei der Untersuchung der *physikalischen Eigenschaften der Minerale und Gesteine* als Grundlage für die Kenntnis deren Genese sind Hochdruckexperimente notwendig, die den betreffenden Tiefenbereich charakterisieren. So gestattet die systematische Untersuchung der Druckabhängigkeit der Schallgeschwindigkeit der verschiedenen Gesteine Rückschlüsse auf Diskontinuitäten in den obersten Erdschichten. Sie ergänzen und interpretieren die durch Bohrungen gewonnenen Resultate. Die exakte Kenntnis der Druck- und Spannungsverhältnisse ist unabdingbar für die Projektierung von Schachtanlagen, Untergrundspeichern, Tunneln, Großbauten und Talsperren sowie deren Überwachung. Die Kenntnis des Druckverhaltens tieferer Erdschichten gibt uns wichtige Rückschlüsse z. B. auf die physikalischen Vorgänge bei *Brucherscheinungen*, die uns wiederum beim Verständnis von Erdbebenvorgängen helfen. Schließlich ist die obere Schicht unserer Erde (Lithosphäre) unser Rohstoff- und Energielieferant. Die exakte Kenntnis der dort ablaufenden Vorgänge in ihrer Komplexität ist nur durch eine systematische Laboruntersuchung unter den extremen Druck- und Temperaturbedingungen zu erhalten. Die *experimentelle Petrologie* untersucht unter diesen Bedingungen einschließlich der notwendigen geochemischen Randbedingungen das Verhalten der Gesteine besonders hinsichtlich ihrer Bildung und Veränderung unter äußeren Einflüssen. Es läßt sich für die gesamte Erde eine *Druck-Tiefen-Charakterisierung* ableiten (Abb. 9.3.13-1), die durch geophysikalische Beobachtungen, experimentelle Laboruntersuchungen und theoretische Bearbeitung ständig verbessert werden wird. Das Ziel der *geophysikalischen Hochdruckforschung* ist neben der Ableitung praktischer Hilfen für die geologisch-geophysikalische Industrie die Schaffung eines detaillierten Bilds über das strukturelle und stoffliche Verhalten unserer Erde. Damit werden die Voraussetzungen geschaffen für eine ge-

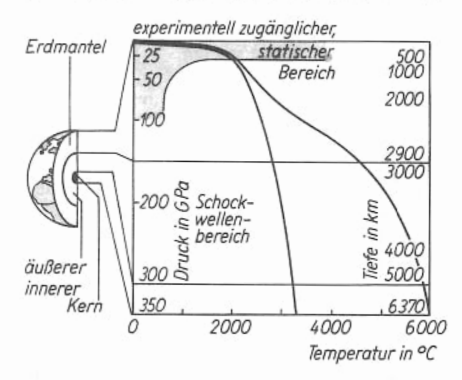


Abb. 9.3.13-1 Druck- und Temperaturverteilung in der Erde

nauere Untersuchung unseres Planetensystems, seines Aufbaus und seiner Entwicklung.
Die *Astrophysik* rechnet bei der Interpretation des stofflichen Verhaltens ihrer Objekte mit dynamischen Drücken in der Größenordnung von Terapascal.

9.4. Gesteuerte Kernfusion

9.4.1. Wesen und Bedeutung der Kernfusion

Angesichts der sich abzeichnenden weltweiten Verknappung fossiler Brennstoffe stellt die Erschließung neuer Energiequellen ein erstrangiges Problem für die Menschheit dar, von dessen Lösung letztlich der weitere technisch-ökonomische Fortschritt abhängen wird. Obwohl vor allem an Steinkohle, in geringerem Maße auch an Erdöl und Erdgas, noch für lange Zeit abbauwürdige Vorräte auf unserem Planeten vorhanden sind, bereitet ihre Erschließung zunehmend höheren Aufwand, da sich viele Lagerstätten in schwer zugänglichen Gebieten, wie den Tundren Sibiriens oder unter dem Meeresboden befinden. Außerdem sind Kohle, Erdöl und Erdgas wichtige Rohstoffe der chemischen Technik, und ihre Verbrennung ist für die weitere Zukunft kaum noch zu verantworten. Als Ausweg für die Energieerzeugung – eigentlich sollte man besser von Energiewandlung sprechen – bieten sich die Kernenergie und die sogenannten alternativen Energiequellen an. Zu den letzteren werden die Solarenergie, Wind- und Wellenenergie sowie die geothermische Energie gerechnet. Sicher werden diese Ressourcen, vor allem unter lokal günstigen Bedingungen, künftig einen wachsenden Beitrag zum Energieaufkommen leisten. Mit Ausnahme der *Solarenergie* (↑ 8.8.3.2.1.), deren Nutzung in großem Maßstab mit Hilfe von irdischen oder kosmischen Stationen jedoch einen extrem hohen Aufwand erfordert, dürften die alternativen Energiequellen langfristig nur eine ergänzende Rolle spielen.
Als dauerhafte Lösung des Energieproblems bleibt somit nur die Erschließung der Kernenergie. Dabei sind zwei Varianten zu unterscheiden:
Die Spaltung oder *Fission* schwerer Atomkerne in Kernreaktoren (↑ 3.4.), die bereits in großem Umfang in Kernkraftwerken realisiert wird. Langfristig ist freilich zu beachten, daß von den in der Natur vorkommenden Nukliden nur U-235, das in gewöhnlichem Uranium lediglich zu reichlich 0,7% enthalten ist, als Kernbrennstoff in Betracht kommt. Weitere Spaltmaterialien, insbesondere Pu-239 oder U-233 müssen erst durch Kernreaktionen aus Uranium oder Thorium „erbrütet" werden, was in gewöhnlichen Kernreaktoren nur sehr unvollkommen möglich ist. Die in Entwicklung befindlichen Brutreaktoren werden die Situation zwar wesentlich verbessern, doch es bleibt die Tatsache, daß Uranium und Thorium recht seltene, noch dazu in der Natur stark verstreute Elemente sind. Schließlich fallen bei der Kernspaltung erhebliche Mengen von radioaktivem Atommüll an, dessen Beseitigung einen hohen Aufwand erfordert.
Die Verschmelzung oder *Fusion* sehr leichter Atomkerne, insbesondere der Wasserstoffnuklide, erfolgt gleichfalls unter Freisetzung von Energie, und die Energiemenge je Masseneinheit Brennstoff ist sogar um etwa eine Größenordnung höher als bei der Kernspaltung. Daß die Kernfusion ebenso wie die Spaltung zur Energiefreisetzung führt, beruht auf der besonderen Stabilität mittelschwerer Atomkerne (↑ 3.2.2.). Die Kernfusion ist auch die wesentliche Energiequelle der Fixsterne (↑ 10.2.). Allein die Sonne strahlt eine Leistung von nahezu 10^{20} kW ab, und der Wasserstoffvorrat reicht noch für mehrere Milliarden Jahre. Unter irdischen Bedingungen wird in erkennbarer Zukunft allerdings nur die Verschmelzung von Kernen der Wasserstoffkerne Deuterium ^2_1H und Tritium ^3_1H, auch als D bzw. T abgekürzt, möglich sein. Während der schwere Wasserstoff D in natürlichem Wasserstoff zu 0,015% enthalten ist und daraus abgetrennt werden kann, muß das radioaktive Tritium T durch Neutronenbestrahlung von Lithium hergestellt werden. Obwohl Lithium auch kein sehr häufiges Element ist, reichen die erschließbaren Vorräte aus, um den erkennbaren Energiebedarf über Jahrhunderte zu decken, und danach wird man sicher die Fusion von reinem Deuterium oder sogar von gewöhnlichem Wasserstoff beherrschen. Selbst bei Beschränkung auf das relativ seltene Deuterium könnte der Energiebedarf der Menschheit auf praktisch unbegrenzte Zeit gedeckt werden, denn der „Heizwert" von dem in 1 Liter Wasser enthaltenen Deuterium entspricht immerhin der bei der Verbrennung von 100 Litern Erdöl anfallenden Energie.
Die vorläufig für die gesteuerte Kernfusion unter irdischen Bedingungen in Betracht kommenden Reaktionen sind in Abb. 9.4.1-1 dargestellt. Die freiwerdende Energie fällt

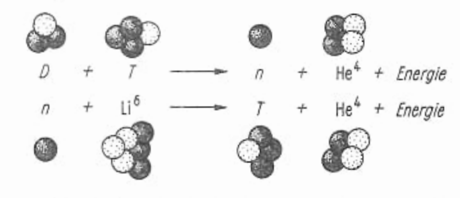


Abb. 9.4.1-1 Wichtige Fusionsreaktionen für die irdische Nutzung

hauptsächlich als kinetische Energie der gebildeten Neutronen (knapp 15 MeV je Teilchen) und in geringerem Maße der Heliumkerne oder α-Teilchen (etwa 3,6 MeV je Teilchen) an. Umgerechnet in konventionelle Energieeinheiten bedeutet dies, daß rund $5 \cdot 10^{18}$ Fusionsprozesse für die Erzeugung 1 kWh Elektroenergie erforderlich sind. Dies entspricht aber nur einer Brennstoffmenge von einigen 10 µg.
In den Fixsternen geht hauptsächlich die Fusion von gewöhnlichem Wasserstoff zu Helium vor sich. Dies erfolgt nach dem von SALPETER angegebenen Mechanismus entweder direkt nach der Reaktionsfolge ${}^1_1H + {}^1_1H \to {}^2_1H + e^+ + \gamma$; ${}^2_1H + {}^1_1H \to {}^3_2He + \gamma$; ${}^3_1He + {}^3_1He \to {}^4_2He + 2{}^1_1H$ oder aber nach dem von BETHE aufgestellten Zyklus unter Beteiligung von Kohlenstoff entsprechend der Folge ${}^{12}_6C + {}^1_1H \to {}^{13}_7N + \gamma$; ${}^{13}_7N \to {}^{13}_6C + e^+ + \gamma$; ${}^{13}_6C + {}^1_1H \to {}^{14}_7N + \gamma$; ${}^{14}_7N + {}^1_1H \to {}^{15}_8O + \gamma$; ${}^{15}_8O \to {}^{15}_7N + e^+ + \gamma$; ${}^{15}_7N + {}^1_1H \to {}^{12}_6C + {}^4_2He$. Man nimmt an, daß die erste Folge bei weniger heißen Sternen, z. B. auch unserer Sonne, den maßgeblichen Beitrag liefert, während der Kohlenstoffzyklus bei heißen Sternen wesentlich beteiligt sein dürfte. Für die irdische Realisierung kommen beide Mechanismen nicht in Betracht, weil sie eine zu große Masse von Brennstoff und zu lange Zeiten benötigen. Bei den Temperaturen von einigen 10^7 K im Inneren der Sterne haben nur wenige besonders energiereiche Kerne eine für die Einleitung des Fusionprozesses ausreichende Energie; jedoch reicht dies angesichts der extrem großen Masse aus, um über astronomische Zeiträume die allmähliche Verschmelzung aufrechtzuerhalten.
Um unter irdischen Bedingungen eine Kernfusion zu „zünden", braucht man sehr viel höhere Temperaturen, als sie im Inneren der Sterne herrschen. Im wesentlichen kommt es darauf an, die beteiligten Kerne auf eine Energie von größenordnungsmäßig 10 keV oder mehr als 10^8 K aufzuheizen, damit sie mit ausreichender Wahrscheinlichkeit die abstoßenden Coulombpotentialwälle überwinden und eine thermonukleare Reaktion eingehen können. Dies war bisher nur bei den Wasserstoffbomben möglich, in denen eine Kernspaltungsbombe als „Zünder" dient, durch den ein Gemisch von wahrscheinlich T + Li oder D + T + Li explosionshaft zur Reaktion gebracht wird.
Für eine friedliche Nutzung der Fusionsenergie müssen Wege gefunden werden, um Brennstoffgemische aus D + T bzw. D + T + Li kontrolliert auf Temperaturen um 10^8 K aufzuheizen und diese Temperatur bei ausreichendem Druck solange aufrechtzuerhalten, daß eine thermonukleare Synthese mit Nettoenergiegewinn gezündet wird. Nach den vor mehr als 30 Jahren begonnenen Pionierarbeiten der von KURTSCHATOW begründeten Forschungsgruppe in der UdSSR erscheinen zwei Varianten als aussichtsreich: 1. das Aufheizen einer Plasmaentladung unter magnetischer Halterung [engl. magnetic confinement] und 2. die Zündung von Brennstoffpillen [engl. pellets] mittels Laser- oder Teilchenstrahlen bei Trägheitshaltung [engl. inertial confinement].
Beide Varianten werden gegenwärtig mit steigendem Aufwand in der UdSSR unter Beteiligung der RGW-Staaten, in den USA, den im Rahmen der EURATOM mitwirkenden westeuropäischen Ländern und in Japan verfolgt. Es ist zu erwarten, daß noch in diesem Jahrzehnt die experimentelle Realisierung von thermonuklearen Reaktionen gelingen wird, die mehr Energie liefern als zur Zündung erforderlich ist. Von dieser wichtigen Etappe bis zur Entwicklung von Fusionskraftwerken wird danach aber noch ein weiter Weg zurückzulegen sein, und man rechnet erst im nächsten Jahrhundert mit einem wesentlichen Beitrag der Kernfusion zur Energiegewinnung.
Ein wichtiges Kriterium für die Beurteilung von Fusionsversuchen bildet der von LAWSON eingeführte Einschlußparameter, nämlich das Produkt $n\tau$ aus der Konzentration der reaktionsfähigen Kerne je Kubikzentimeter und der Zeitspanne τ während der das aufgeheizte Plasma gehaltert werden kann. Nach dem *Lawsonkriterium* muß das Produkt wenigstens $6 \cdot 10^{13}\ cm^{-3}s^{-1}$ betragen, damit sich eine thermonukleare Reaktion energetisch selbst tragen kann, jedoch sind für einen brauchbaren Nettoenergiegewinn höhere Werte von etwa $5 \cdot 10^{-14}\ cm^{-3}s^{-1}$ anzustreben.
Zur Erreichung dieser Ziele ist eine Beherrschung von extremen Stoffzuständen notwendig, wie sie bisher nur aus kosmischen Verhältnissen bekannt sind. Zu den wissenschaftlich-technischen Voraussetzungen gehören die neuesten Erkenntnisse der Plasmaphysik und der Atomkernphysik ebenso wie Anwendungen der Hochvakuum- und Kryotechnik sowie der Festkörperphysik in bisher nicht gekanntem Maßstab. Man darf erwarten, daß auf dem Weg zur Erschließung der kontrollierten Kernfusion ein Innovationsschub an neuen Erkenntnissen einsetzt, der sich stimulierend auf weite Gebiete des wissenschaftlich-technischen Fortschritts auswirken wird.

9.4.2. Konzepte der gesteuerten Kernfusion mit magnetischer Halterung

9.4.2.1. Prinzipien der magnetischen Halterung. Die Grundidee dieser Variante besteht darin, das Plasma einer in Mischungen von Deuterium und Tritium unterhaltenen Gasentladung (↑ 9.2.3.) durch ohmsche Aufheizung und Teilcheneinschuß auf höchste Temperaturen zu

bringen und dieses heiße Plasma von den Gefäßwänden durch starke Magnetfelder zu isolieren. Dazu wurden verschiedene Arten magnetischer Fallen entwickelt.
Bei den *linearen Magnetfallen* (Abb. 9.4.2-1), auch offene oder Spiegelmagnetfallen genannt,

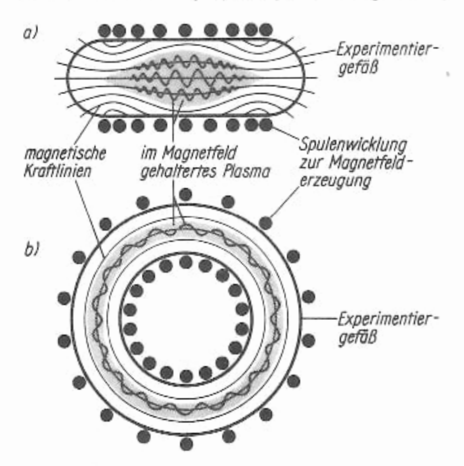


Abb. 9.4.2-1 Prinzip von Magnetfallen; a) lineare, b) toroidale Anordnung

ist das spindelförmige Plasma durch Magnetfeldlinien eingeschlossen, die sich an den Enden des Gefäßes verdichten. An den Gefäßenden bilden sich auf diese Weise magnetische Spiegel aus, die ankommende geladene Teilchen reflektieren. Auch von den Gefäßwänden werden die Plasmateilchen weitgehend ferngehalten, weil sie die longitudinalen Magnetfeldlinien kreuzen müßten, um die Wände zu erreichen. Bekanntlich können Ionen oder Elektronen eines Plasmas Magnetfeldlinien ausreichender Stärke aber nicht kreuzen, sondern sie werden so abgelenkt, daß sie die Feldlinien auf schraubenförmigen Bahnen umlaufen (↑ 9.1.3.5.). Nach vielen Jahren grundlegender Untersuchungen, bei denen die Beherrschung der zahlreichen Plasmainstabilitäten gelöst werden mußte, gelang es, ein Plasma von 10^{14} Teilchen je Kubikzentimeter auf etwa 10^8 K aufzuheizen. Dies erfolgte vor allem durch Einschuß intensiver neutraler Strahlen von Wasserstoffatomen hoher kinetischer Energie. Leider beträgt die Einschlußzeit τ nur einige Mikrosekunden, weil zu viele Plasmateilchen durch die „löchrigen" Magnetspiegel entkommen. An Verbesserungen dieser Anlagen wird intensiv gearbeitet. Aussichtsreicher erscheinen die ringförmig geschlossenen Plasmafallen, in denen die Ionen Umlaufbahnen beschreiben. Von den verschiedenen Ausführungsformen hat sich der *Tokamaktyp* durchgesetzt, der zuerst in der UdSSR entwickelt worden ist. Abb. 9.4.2-2 und Abb. 9.4.2-3 zeigen den prinzipiellen Aufbau eines Tokamaks und die Ausführung der ringförmigen Kammer. Die magnetische Halterung geschieht durch toroidale und poloidale Magnetfelder, die von leistungsstarken Feldspulen passender Anordnung erzeugt werden. Um den gewaltigen Energieverbrauch dieser Spulen einzuschränken und zugleich den in hunderte von Tonnen gehenden Kupfereinsatz herabzusetzen, werden gegenwärtig für einige Tokamakanlagen supraleitende Magnetspulen erprobt bzw. projektiert (↑ 8.5.4.). Die Entladungskammern müssen vor den Versuchen mit leistungsfähigen Pumpsystemen auf Ultrahochvakuum ausgepumpt werden. Man benutzt hierfür zunehmend Kryopumpen, die mit flüssigem Helium beschickt werden (↑ 9.2.2.). Die Tatsache, daß für die Erzeugung höchster Plasmatemperaturen Hilfsmittel der extremen Tieftemperaturphysik in bisher nicht gekanntem Maßstab eingesetzt werden, unterstreicht die Komplexität von Versuchsanlagen zur Erschließung der gesteuerten Kernfusion.
Zum Betrieb eines Tokamaks läßt man zunächst allmählich Gas ein (für die Erprobung werden gewöhnlich nur H und D eingesetzt) und zündet durch Einspeisen von Hochfrequenzenergie eine intensive Gasentladung. Unter der Wirkung des eigenen vom Entladungsstrom aufgebauten Magnetfelds kommt es zum Einschnüren der Entladung, zum *Pincheffekt* [to pinch engl., einschnüren]. Dieser Vorgang wird durch eine entsprechende Steuerung der äußeren Spulenfelder verstärkt. Bei andauerndem Gaseinlaß und Erhöhung des Entladungsstroms erfolgt in der fadenförmigen Entladung eine intensive Ohmsche Aufheizung. Bereits mit dem Tokamak T-3 der 60er Jahre, der von ARTSIMOWITSCH am Kurtschatow-Institut entwickelt wurde, konnte die Ionentemperatur im Plasma auf 0,5 keV oder auf mehrere 10^6 K gesteigert werden,

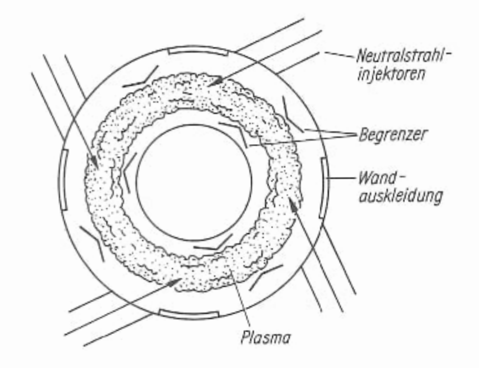


Abb. 9.4.2-2 Aufbau eines Tokamakreaktors (schematisch)

jedoch wurde wegen Plasmainstabilitäten nur eine Einschlußzeit von 5 bis 10 ms erreicht. Dieser Tokamak besaß einen großen Radius von 100 cm und einen kleinen Radius von 15 cm. Das toroidale Magnetfeld betrug 65000 Gauß ≙ 0,065 Tesla und der Entladungsstrom erreichte 10^5 A. Inzwischen sind mindestens 50 Tokamaks gebaut worden, darunter solche mit erheblich größeren Dimensionen und gesteigerten Leistungsparametern. So wurden in der sowjetischen Anlage T-10 und der ähnlichen amerikanischen Anlage PLT Temperaturen von 10^7 K bei Einschlußzeiten von knapp 0,1 s und einer Teilchendichte von $10^{14}\ \mathrm{cm}^{-3}$ erreicht. Um das Lawsonkriterium (↑ 9.4.1.) zu erfüllen, müssen also sowohl die Temperatur als auch die Einschlußzeit um eine weitere Größenordnung gesteigert werden. Die hierfür einzuschlagenden Wege sind im Prinzip klar. Erstens muß man die Abmessungen der Anlagen erhöhen, etwa auf große Torusradien von 3 bis 5 m und kleine Radien von 1 m, was den ohnehin schon großen Aufwand natürlich außerordentlich steigert. Zweitens sind nach dem Vorbild der mit Spiegelfallen erzielten Resultate zusätzliche Aufheizverfahren, besonders die Injektion neutraler Wasserstoffstrahlen, zu entwickeln. Entsprechende Tokamaks dieser neuen Generation sind in der UdSSR (T-15, T-20), den USA (TFTR), Japan (JT-60) und anderen Ländern im Bau, und man darf erwarten, daß in einigen Jahren der experimentelle Nachweis einer sich selbst tragenden gesteuerten Kernfusion gelingt.

Im übrigen sind schon bei den bisher durchgeführten Fusionsversuchen Kernverschmelzungsreaktionen nachgewiesen worden. Die dabei freigesetzte Energie war jedoch viel kleiner als der zur Einleitung erforderliche Aufwand.

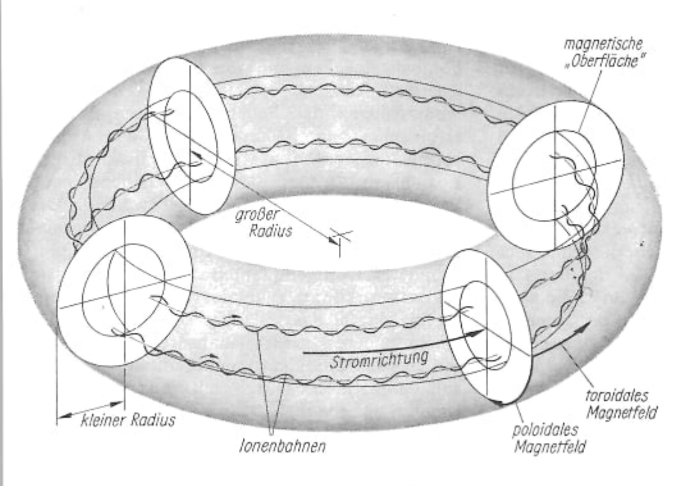


Abb. 9.4.2-3 Feldverlauf und Teilchenbahnen im Tokamak

9.4.2.2. Plasma-Wand-Wechselwirkung. Bei der Entwicklung verbesserter Fusionsversuchsanlagen mit magnetischer Halterung sind naturgemäß zahlreiche Probleme zu lösen, die viele Gebiete der Physik und Technik berühren. Herausgegriffen sei die besonders wichtige Problematik der Plasma-Wand-Wechselwirkung und der damit zusammenhängenden Effekte. Die Wände der Entladungskammern sind stets einem Bombardement von energiereichen Teilchen ausgesetzt, die aus der Magnethalterung entkommen, und nach Ingangsetzen der Fusionsreaktion treten die energiereichen Neutronen von reichlich 14 keV hinzu, die als neutrale Teilchen die Magnetfelder mühelos durchqueren können. Beide Erscheinungen führen zu einer allmählichen Erosion der Wände durch Ionenzerstäubung sowie zu schädlichen strukturellen Veränderungen. Am schlimmsten wirkt sich die Tatsache aus, daß mit dem abgestäubten Material schwere Atome in das Plasma gelangen. Bereits kleine Verunreinigungen führen aber zu einer starken Abkühlung, die hauptsächlich durch Energieverluste infolge Emission von Röntgenstrahlen verursacht wird. Aus diesem Grunde ist auch das Auspumpen auf Ultrahochvakuum zur Entfernung von Luftbestandteilen und die Verwendung reinster Gase unumgänglich. Die wichtigsten Maßnahmen zur Unterdrückung der schädlichen Plasma-Wand-Wechselwirkung bestehen in der Beschichtung der Wände mit schwer zerstäubbaren Materialien niederer Ordnungszahl, wie C, B oder Borcarbid, und im Anbringen weiterer Magnetfelder, um wenigstens die Ionen von den Wänden fernzuhalten. An der Lösung dieses keineswegs nebensächlichen Problems wird intensiv gearbeitet.

9.4.2.3. Erkennbare Trends. Auch nach erfolg-

reicher Realisierung der laufenden Projekte wird man noch nicht daran gehen können, den Bau von Versuchskraftwerken in Angriff zu nehmen. Als Zwischenetappe ist zweifellos der Aufbau einer Demonstrationsanlage notwendig, in der das Zusammenspiel aller wesentlichen Systeme unter den späteren Bedingungen einer Fusionsreaktion im Langzeitbetrieb getestet werden muß. Die UdSSR hat vorgeschlagen, eine solche Demonstrationsanlage INTOR, für die ein Kostenaufwand in Höhe von mindestens einer Milliarde Rubel veranschlagt wird, in internationaler Kooperation zu errichten und zu betreiben. Dieser Vorschlag hat ein positives Echo in vielen Ländern gefunden, und die Realisierung könnte etwa bis zum Jahre 1990 erfolgen. Danach ist die Entwicklung erster Versuchskraftwerke zu erwarten, die frühestens am Anfang des kommenden Jahrhunderts in Betrieb gehen könnten, falls es gelingt, die noch weitgehend ungeklärten Probleme der Energieabfuhr zu lösen. Als aussichtsreiche Übergangslösung wird der Bau von *Fusionshybridreaktoren* angesehen, in denen die bei der Fusion gebildeten schnellen Neutronen genutzt werden, um aus Uranium das für Kernspaltungskraftwerke benötigte Plutonium herzustellen. Diese Variante stellt eine interessante und wahrscheinlich technisch und ökonomisch günstige Alternative zum schnellen Brutreaktor dar.

9.4.3. Konzepte der gesteuerten Kernfusion mit Trägheitshalterung

9.4.3.1. Prinzip der Trägheitshalterung. In den letzten Jahren ist ein anderer Weg zur kontrollierten Durchführung thermonuklearer Reaktionen beschritten worden. Das Prinzip dieser Kernfusion mit Trägheitshalterung ist in Abb. 9.4.3-1 dargestellt. Das Brennstoffgemisch, verflüssigtes D + T, ist in einer kleinen Glaskugel von 0,1 bis 1 mm innerem Durchmesser eingeschlossen. Diese Kugel ist mit mehreren dicken Schichten, z. B. einer inneren Polymerschicht und einer äußeren Metallschicht überzogen. Der Durchmesser dieser Pillen oder Pellets beträgt insgesamt höchstens wenige Millimeter. Auf ein solches Objekt werden nun gleichzeitig von allen Seiten extrem leistungsstarke Impulse von Laser-, Elektronen- oder Ionenstrahlen gerichtet. Dadurch wird innerhalb weniger milliardstel Sekunden eine enorme Aufheizung bewirkt. Als Folge verdampft das äußere Umhüllungsmaterial explosionsartig. Dabei wird ein Rückstoß erzeugt, der zusammen mit dem Strahlungsdruck das innere Verkleidungsmaterial nach innen treibt. Die extreme Kompression des Brennstoffs auf die mehrhundertfache Dichte des kondensierten Materials bei gewöhnlichem Druck ist von einer gewaltigen Aufheizung begleitet. Die Theorie derartiger Schockwellenprozesse ist kompliziert. Nach den bisherigen experimentellen Ergebnissen und theoretischen Untersuchungen erscheint es möglich, Temperaturen von etwa $3 \cdot 10^8$ K zu erreichen. Unter diesen Bedingungen müßte der Brennstoff innerhalb einer milliardstel Sekunde eine thermonukleare Reaktion nach Art einer „Miniwasserstoffbombe" eingehen, die von einer schlagartigen Energiefreisetzung begleitet ist. Die kleinen Abmessungen der Pellets sind auch deshalb erforderlich, damit Zerstörungen der Anlage vermieden werden. Dafür sollen in künftigen Fusionskraftwerken dieser Art laufend neue Pellets zugeführt und gezündet werden. Die Energie wird vorwiegend in Form von energiereichen Neutronen, Ionen und auch sekundär gebildeten Röntgenquanten abgegeben. Um sie zu nutzen, könnte man das Reaktorgefäß mit einem Mantel oder Blanket aus Lithium oder einem anderen neutronenabsorbierenden Metall, das dabei möglichst nicht radioaktiv werden soll, auskleiden. Durch ein Rohrsystem im Inneren des Blankets ließe sich die entstandene Wärme mit geeigneten Kühlmitteln abführen und in Wärmeaustauschern zur Erzeugung von Dampf nutzen.

9.4.3.2. Stand und Trends. Obwohl die Kernfusion mit Trägheitshalterung erst später und zunächst mit geringerem Aufwand als die Variante mit magnetischer Halterung untersucht worden ist, sind in den letzten Jahren Großanlagen entstanden bzw. im Bau, und die

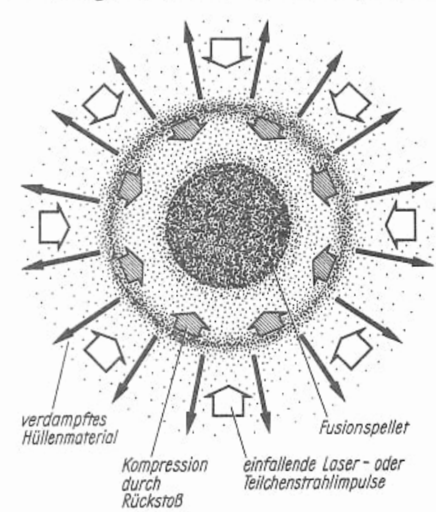


Abb. 9.4.3-1 Prinzip der gesteuerten Kernfusion mittels Inertial- bzw. Trägheitshalterung; das Fusionspellet ist ein mit D + T gefülltes Glaskügelchen

ersten ermutigenden Versuchsergebnisse haben eine Beschleunigung und Aufstockung der Projekte nach sich gezogen. Ebenso wie bei der Tokamakvariante ist noch kein Durchbruch im Sinne einer Fusionsreaktion geglückt, die mehr Energie freisetzt als zur Einleitung verbraucht wird. Immerhin konnten jedoch Fusionsprozesse nachgewiesen werden, und die eigentliche Zündung der Pellets dürfte im kommenden Jahrzehnt realisierbar sein. Der Aufwand ist freilich außerordentlich hoch. Bei den Anlagen Delphin in der UdSSR und Shiva-Nova in den USA arbeitet man mit einer größeren Zahl von Hochleistungslasern (20 bis 30), insbesondere CO_2- oder N_2-Gaslasern von vielen Metern Länge. Über gewaltige Kondensatorbatterien und eine präzise Ansteuerelektronik werden die Laser gleichzeitig im Impulsbetrieb gezündet. Das ausgestrahlte Licht wird über eine hochkomplizierte Optik so abgelenkt und fokussiert, daß das frei fallende Pellet allseitig von der Strahlung getroffen wird. Nach neueren Meldungen scheint allerdings gerade eine definierte Unsymmetrie durch Superposition der Schockwellen günstig zu sein. Die Strahlungsleistung, die während der extrem kurzen Impulsdauer auf die Pellets von einigen Millimetern Durchmesser einwirkt, ist dabei größer als die Leistung der gesamten gegenwärtigen Erzeugung von Elektroenergie auf der Erde! Das unterstreicht erneut, daß die Erschließung der Fusionsenergie auf eine Technik in neuen Dimensionen hinausläuft (Tafel 49).

Ähnliche Maßstäbe gelten für Anlagen, bei denen man anstelle der Laserstrahlen intensive Teilchenstrahlimpulse von Elektronen oder Ionen einsetzt. Eine Anlage, die mit Elektronenstrahlen arbeitet, wird unter dem Namen Angara-5 in der UdSSR aufgebaut; die Entwicklung von Ionen- oder Neutralstrahlanlagen wird im sogenannten Sandia-Projekt in den USA verfolgt.

Parallel zur Entwicklung der Laser- und Teilchenstrahlquellen nebst der zugehörigen Energieversorgung ist eine Fülle von weiteren Problemen zu bearbeiten. Genannt seien die noch mangelhaft gelöste Entwicklung von Vergütungsschichten (↑ 8.8.1.5.) für die stark belasteten optischen Systeme der Laseranlagen, die optimale Fokussierung und Einspeisung der intensiven Teilchenstrahlen sowie die wahrscheinlich notwendige Abschwächung der Explosionsstoßwelle durch geeignete Restgase im Reaktionsgefäß. Eine vergleichende Einschätzung der Perspektiven von Fusionslösungen auf der Basis der magnetischen Halterung und solchen mit Trägheitshalterung fällt gegenwärtig schwer. Zweifellos ist die Tokamakvariante zur Zeit am weitesten fortgeschritten. Die Anhänger der Laser- und Teilchenstrahlkonzeptionen führen jedoch zu Gunsten dieser Lösungen an, daß die Weiterentwicklung zum Kraftwerk unter Nutzung konventioneller Dampfturbinen leichter und schneller möglich sein sollte als beim Tokamak. Im übrigen lassen sich auch mit Fusionsreaktoren vom Typ der Trägheitshalterung Hybridvarianten ausdenken, bei denen die freigesetzten Neutronen zur Herstellung von Spaltmaterial für Kernspaltungskraftwerke genutzt werden.

10. Kosmische Erscheinungsformen der Materie

10.1. Planet Erde und erdnaher Raum

Durch Untersuchung des radioaktiven Zerfalls (↑ 3.4.3.) wurde festgestellt, daß die ältesten Gesteine der Erde vor etwa $4 \cdot 10^9$ Jahren entstanden sind, die Entstehung der Erde selbst also noch früher erfolgte. Der Erdkörper (s. a. 10.1.1.) ist von einer Atmosphäre umgeben (↑ 10.1.2.), die mit zunehmender Höhe unter der Wirkung ionisierender Strahlen und des erdmagnetischen Felds in die Ionosphäre (↑ 10.1.3.) und die Magnetosphäre (↑ 10.1.4.) übergeht. Die Atmosphäre absorbiert die von der Sonne kommende Wellenstrahlung bis auf das sichtbare Licht und einen Teil der Strahlung im Radiowellenbereich. Das Magnetfeld der Erde wirkt als Schutzschild gegen die von der Sonne als Sonnenwind (↑ 10.1.4.1.) und aus dem Weltall einfallenden hochenergetischen Teilchen. Atmosphäre und Erdmagnetfeld stellen damit eine wichtige Voraussetzung für die Entstehung und Erhaltung des Lebens auf der Erde dar.

10.1.1. Erdkörper

Die Erde ist ein durch ihre Masseverteilung bestimmtes Geoid, das in 0. Näherung als Kugel von einem mittleren Radius von 6371,2 km und 1. Näherung als Rotationsellipsoid mit dem Äquatorhalbmesser $a = 6378{,}388$ km und dem Polhalbmesser $c = 6356{,}909$ km angesehen wird. Die Abplattung $A = (a - c)/a$ beträgt 1/293,25, die Masse $5{,}98 \cdot 10^{23}$ kg. 70% der Oberfläche sind mit Wasser bedeckt, das eine Masse von $1{,}43 \cdot 10^{21}$ kg hat und eine mittlere Tiefe von 3792 m.

10.1.1.1. Physikalisch-strukturelles Bild des Erdkörpers. Aussagen über das Erdinnere sind nur indirekt durch die Untersuchung der *Ausbreitung von Erdbebenwellen* möglich. Neben den das Erdinnere durchlaufenden Raumwellen, zu denen die schnelleren Verdichtungswellen oder P-Wellen [*prima unda* lat., erste Welle] sowie die langsameren Scherungs- oder S-Wellen [*secunda unda* lat., zweite Welle] gehören, werden in letzter Zeit auch Oberflächenwellen und Eigenschwingungen des Erdkörpers zur Untersuchung des Erdinneren herangezogen. Oberflächenwellen sind Wellen, die an der Grenzfläche zwischen zwei Körpern mit verschiedenen physikalischen Eigenschaften geführt werden und neben gebrochenen und reflektierten Wellen beim Auftreffen von Raumwellen auf eine derartige Grenzschicht entstehen. Zur Deutung der Beobachtungen werden theoretische Modelle für die Wellenausbreitung im Erdinneren aufgebaut und den Beobachtungen angepaßt. In diesen Modellen spielen Zustandsgleichungen für den Zusammenhang von Druck und Dichte in verschiedenen Tiefen eine Schlüsselrolle. Kennt man diese, so kann man auf indirektem Wege Schlüsse über die Eigenschaften der Materie in verschiedenen Schichten des Erdinneren ziehen.

10.1.1.1.1. *Elastomechanische Eigenschaften.* Aus den Reaktionen auf Erdbeben und die Wirkung der Erdrotation ergibt sich, daß sich der Erdkörper im Mittel wie eine ideale Flüssigkeit verhält und eine hydrostatische Gleichgewichtsfigur darstellt. Die Untersuchung der Geschwindigkeit der Erdbebenwellen in Abhängigkeit von der Tiefe hat zu einem *vertikalen Schichtenaufbau des Erdinneren* geführt (Abb. 10.1.1-1). Für die einzelnen, durch ihre Tiefenlage in Kilometer gekennzeichneten Schichten sind folgende Bezeichnungen und Abkürzungen üblich: Erdkruste A bis 33 km; oberer Erdmantel B bis 410 km und C bis 1000 km; unterer Erdmantel D' bis 2700 km und D'' bis 2900 km; äußerer Erdkern E bis 4980 km; Übergangsbereich F bis 5120 km; innerer Erdkern G bis 6371 km.

Die Erdkruste hat in verschiedenen Teilen der Erde verschiedene Dicke. Im allgemeinen folgt auf eine Sediment- eine Granit-Gneis-Schicht und anschließend die Gabbroschicht. Der Krustenbereich wird durch die *Mohorovičicdiskon-*

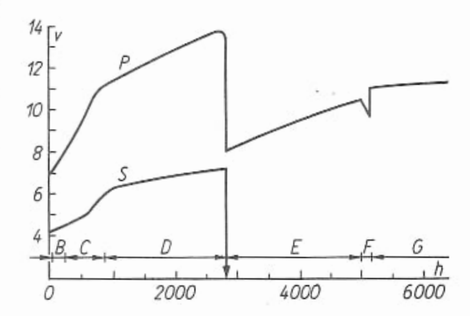


Abb. 10.1.1-1 Geschwindigkeit v in km s⁻¹ seismischer Verdichtungs- (P-) und Scherungs- (S-) Wellen in Abhängigkeit von der Tiefe h in km

tinuität abgeschlossen. Diese stellt eine rein mechanische Diskontinuität dar, während andere Diskontinuitäten häufig mit sprunghaften Änderungen anderer physikalischer Parameter verbunden sind. Im äußeren Teil des oberen Mantels beginnt ein starker Geschwindigkeitsanstieg, der dann in die sogenannte Zone niedriger Geschwindigkeiten nach GUTENBERG übergeht. Im oberen Mantel treten bei etwa 400 und 700 km weitere Diskontinuitäten auf, deren Tiefenlage regional um 100 bis 200 km variiert. Im unteren Erdmantel, der sich bis 2900 km Tiefe erstreckt, treten keine weiteren Diskontinuitäten auf. Der Übergang vom Erdmantel zum Erdkern ist durch eine starke Diskontinuität und das Verschwinden der Scherungswellen gekennzeichnet, aus dem man schließt, daß der äußere Erdkern eine hochkomprimierte Flüssigkeit darstellt. Zwischen äußerem und innerem Erdkern wird ein Übergangsbereich beobachtet. Am Rand des inneren Erdkerns erfolgt ein starker Geschwindigkeitsanstieg, der gegenwärtig nur durch eine erhöhte Scherfestigkeit gedeutet werden kann, so daß man zu der Annahme gezwungen ist, daß der innere Erdkern fest sei.
Die Abhängigkeit der Dichte, des Drucks und der Schwerebeschleunigung von der Tiefe ist in Tab. 10.1.1-2 zusammengefaßt. Da die Kompressibilität im Erdinneren nicht ausreicht, um die beobachteten Dichtezunahmen zu erklären, muß das Erdinnere in bestimmten Schichtbereichen inhomogen sein. Es muß also verschiedene Phasen bzw. Phasenmodifikationen der Materie im Erdinneren geben.

10.1.1.1.2. *Thermische, elektrische und viskose Eigenschaften.* Man kann heute die thermischen Verhältnisse im Erdinneren nur grob abschätzen. Da alle Wärmetransportvorgänge im Erdinneren außerordentlich langsam ablaufen, muß davon ausgegangen werden, daß z. B. für die Temperaturverteilung auch heute noch die Temperaturverteilung bei der Entstehung des Planeten mitbestimmend ist und daß man zum Verständnis der geothermischen Verhältnisse die gesamte thermische Geschichte der Erde kennen muß. Alle bisher vorgelegten Hypothesen zur Entstehung und Entwicklung der Erde zeigen einheitlich, daß sich das Innere der Erde ständig in einem überhitzten Zustand befunden hat und noch befindet. Zu allen Zeiten ist die Wärmeproduktion im Erdinneren größer gewesen als die Wärmeabgabe durch die Erdoberfläche.

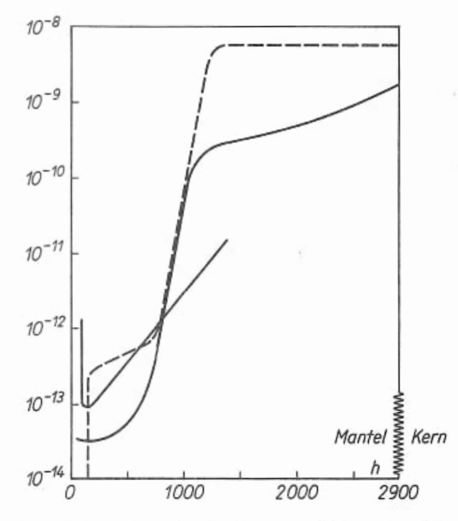


Abb. 10.1.1-3 Elektrische Leitfähigkeit im Erdmantel in elektromagnetischen Einheiten; *h* Tiefe in km

Die *Temperaturverteilung im Erdinneren* kann nur indirekt durch Abschätzungen erschlossen werden. Als Quellen des Wärmestroms werden der radioaktive Zerfall langlebiger Isotope des Uraniums, Thoriums und Kaliums in der Erdkruste und im Erdmantel sowie weitere mögliche wärmeerzeugende Prozesse vermutet. Als Wärmetransportprozesse berücksichtigt man in den festen Bereichen des Erdinneren Wärmeleitung und in größeren Tiefen Wärmestrahlung sowie im oberen Erdmantel durch Kriechpro-

Tab. 10.1.1-2 Elasto-mechanische Größen im Erdinneren (Erläuterung im Text)

Tiefe in km	Lage	Dichte in g cm^{-3}	Druck in 10^{11} Nm^{-2}	Schwere in cm s^{-2}	Inkompressibilität in 10^{11} N m^{-2}	Righeit in 10^{11} N m^{-2}	Poissonverhaltnis
35	Beginn *B*	3,3	0,01	985	1,2	0,6	0,27
1000	Grenze *C*/*D*	4,7	0,4	995	3,6	1,9	0,28
2900	Ende *D*	5,7	1,3	1030	6,5	3,0	0,38
2900	Beginn *E*	9,7	1,3	1030	6,2	0,0	0,50
5000	in *F*	(12,5)	3,2	(500)	(12,8)	(0,2)	(0,49)
6370	*G*	(13,0)	3,7	0	(14,7)	(1,3)	(0,46)

zesse realisierte Konvektion. In den flüssigen Bereichen ist Konvektion der vorherrschende Prozeß. Der Wärmestrom durch die Erdoberfläche mit $6{,}3 \cdot 10^{-2}\ \mathrm{Jm^{-2}s^{-1}}$ als äußere Randbedingung ist relativ gut bekannt. Die Temperatur ändert sich mit der Tiefe sehr viel stärker als der Druck. Höchstwahrscheinlich sind die Wärmequellen im Erdinneren gleichzeitig die Energiequellen für alle anderen Bewegungsvorgänge und spielen so eine Schlüsselrolle.
Die *elektrische Leitfähigkeit* in Abhängigkeit von der Tiefe ist in Abb. 10.1.1-3 durch verschiedene Modellkurven dargestellt. Ein Trend zunehmender Leitfähigkeit, d. h. abnehmenden Widerstands mit zunehmender Tiefe, ist deutlich erkennbar. Es scheint sich anzudeuten, daß zum unteren Erdmantel hin Halbleitereigenschaften überwiegen. Extrapoliert man die Verhältnisse zum äußeren Erdkern hin, so muß man, ausgehend von Hochdruckmeßergebnissen, metallische Leitfähigkeitverhältnisse vermuten.
Die *Rheologie des Erdinneren*, d. h. die Untersuchung der Fließerscheinungen der flüssigen und festen Bereiche des Erdinneren, wird z. Z. sehr intensiv betrieben. Rheologische Parameter, vor allem die Viskosität, werden für das Verständnis aller dynamischen Prozesse im Erdinneren, z. B. großräumiger Bewegungen im Erdmantel wie Kontinentaldriften und Erdbeben, oder von Bewegungsvorgängen im flüssigen äußeren Erdkern dringend benötigt. Die Viskosität im oberen Erdmantel ist mit 10^{20} Pas recht gering. Für den unteren Erdmantel streuen die Werte der bisherigen Abschätzungen von 10^{22} bis 10^{25} Pas. Für den flüssigen äußeren Erdkern werden Werte in der Größenordnung 10^{2} bis 10^{6} Pas angegeben. Für den inneren Erdkern erreichen die Schätzwerte die gleiche Größenordnung wie im unteren Erdmantel, so daß auch aufgrund dieser Analysen die Existenz eines festen Erdkerns plausibel erscheint.

10.1.1.2. Materialeigenschaften im Erdinneren. Um die *Tiefenabhängigkeit* der verschiedenen physikalischen Parameter zu verstehen, ist es erforderlich, die speziellen stofflichen Zusammensetzungen und Strukturen der einzelnen Bereiche zu erforschen. Aus Tiefenbohrungen bis 10 km und durch Analyse von Hebungen und Faltungen, die Schichten bis 15 km Tiefe erschließen, kennt man die neun häufigsten Elemente in der Erdkruste, die zusammen 99% der Masse ergeben: O 48,3%; Si 29,7%; Al 8,4%; Fe 4,2%; K 2,9%; Ca 2,3%; Na 1,8%; Mg 1,5%; Ti 0,5%. Die oberste Schicht der Sedimente wird an zahlreichen Stellen von der Granit-Gneis-Schicht durchstoßen.
Für das Verständnis der *stofflichen Zusammensetzung* in tieferen Schichten spielen indirekte Verfahren und experimentelle Laboruntersuchungen eine große Rolle, bei denen verschiedenartige Materialien unter mit dem Erdinneren vergleichbaren Druck- und Temperaturbedingungen untersucht werden. Statische Druckanlagen werden in den nächsten Jahren im Labor Drücke erreichen, wie sie im unteren Erdmantel herrschen dürften. Höhere Drücke sind nur durch Stoßwellenexperimente erzeugbar (↑ 9.3.). Die Verkopplung der Laborergebnisse und geophysikalischer Messungen erfolgt über Zustandsgleichungen, die die Umrechnung der im Labor erzielten Ergebnisse in die verschiedenen Tiefenbereiche im Erdinneren und umgekehrt die Reduktion der geophysikalischen Meßdaten auf die Laborverhältnisse gestatten.
Die Hochdruckexperimente und die geophysikalischen Resultate ergeben folgendes Bild der stofflichen Verhältnisse mit zunehmender Tiefe: Die untere Erdkruste besteht vorrangig aus Normaldruck-Alumosilicaten, die unter der Wirkung des steigenden Drucks dann in Hochdrucksilicate übergehen. Bei etwa 400 km Tiefe herrschen Druck- und Temperaturverhältnisse, in denen Hochdrucksilicate instabil werden und sich in eine stabile Spinellstruktur umwandeln. In etwa 700 km Tiefe ist der Druck so angestiegen, daß auch die Spinellstrukturen nicht mehr stabil sind. Die komplexen Silicate zerfallen und gehen in einfache Oxide über, deren Eigenschaften den gefundenen elastomechanischen Werten entsprechen. Unter der Wirkung des zunehmenden Drucks wandeln sich die einfachen Oxide in Hochdruckoxide um, die für den ganzen unteren Erdmantel charakteristisch sind und den gefundenen Dichtewerten von $6\ \mathrm{g\,cm^{-3}}$ gut entsprechen.
Es liegen viele Anzeichen dafür vor, daß äußerer und innerer Erdkern aus Mischungen von etwa 80% Eisennickel und 20% leichteren Elementen wie Silicium, Schwefel, Sauerstoff bestehen. Das würde dem vermuteten Dichtewert von $9\ \mathrm{g\,cm^{-3}}$ entsprechen.
Verbunden mit der stofflichen Änderung der Materialien im Erdinneren ist auch eine *Änderung des Bindungsverhaltens*. Bei den Silicaten im oberen Erdmantel überwiegt Ionenbindung. Die gefundenen Werte der elektrischen Leitfähigkeit entsprechen den bei Ionenbindung zu erwartenden Werten. Unter der Wirkung des zunehmenden Drucks im oberen Erdmantel verringern sich die Abstände der wechselwirkenden Atome. Es stellen sich zunehmend kovalente Bindungen ein. Ein sukzessiver Übergang zu kovalenter Bindung erfolgt im unteren Erdmantel, d. h. ein Übergang zu mehr metallischen Bindungsanteilen mit zunehmender Tiefe. Sowohl der zunehmende Metallisierungsgrad als auch das elektrisch mit Halbleiterleitfähigkeit vergleichbare kovalente Bindungsverhalten entsprechen den geophysikalischen Beobachtungen im unteren Erdmantel. Für den liquiden äußeren Erdkern lassen Laborexperimente und Dynamotheorie (↑ 10.1.1.4.) metallische Leit-

fähigkeit und metallisches Bindungsverhalten erwarten.

10.1.1.3. Dynamische Vorgänge in den äußeren Bereichen. Abweichungen vom Bild der Erde als Drehkörper einer idealen Flüssigkeit lassen sich bis zu 700 km Tiefe nachweisen und bis 1000 km Tiefe vermuten. Beachtet man interne dynamische Vorgänge und Bewegungen im Erdinneren, so macht sich eine andere vor allem die Festigkeit beachtende Gliederung der oberen Schichten erforderlich. Man unterscheidet die starre, spröde *Lithosphäre* (bis 100 km), die zähflüssige *Asthenosphäre* (100 bis 200 km) und die schwach fließfähige *Mesosphäre* (200 bis 700 km). Die Lithosphäre, die mit der Gutenbergschen Zone endet, ist horizontal stark gegliedert in in sich feste und stabile Platten, die sich parallel zur Erdoberfläche bewegen. In den Riftzonen entfernen sie sich voneinander, z. B. Amerika von Europa längs der Riftzone im Atlantik. In Kollisionszonen taucht die eine unter die andere und ruft starke tektonische, seismische und vulkanologische Erscheinungen hervor. Die Asthenosphäre, die die Gutenbergsche Zone umfaßt, ist durch geringe Dichte und Viskositätswerte gekennzeichnet. Aufgrund der geringen Dichte dürfte sie eine Quelle mechanischer Instabilität sein. Die geringe Viskosität sowie andere Ergebnisse deuten darauf hin, daß das Material in ihr teilweise geschmolzen ist. Durch diese zähflüssige Schicht wird die Lithosphäre recht wirkungsvoll von der Mesosphäre und den unteren Teilen des Erdmantels entkoppelt. Tangentiale Spannungen können wegen der geringen Viskosität der Asthenosphäre nicht direkt auf die Lithosphäre übertragen werden. Eine mechanische Kopplung erfolgt daher lediglich in Gebieten, wo die Lithosphäre direkt aus der Mesosphäre aufsteigt oder wo sie im Mantel verschluckt wird, z. B. dort, wo Platten übereinander gleiten. Da die Plattenbewegungen große Massenverlagerungen darstellen, erfordern sie wegen der Erhaltung der Masse ebenso große Strömungen in entgegengesetzter Richtung im Erdinneren, also eine Konvektion im Erdmantel.

10.1.1.4. Erdmagnetisches Hauptfeld. Die Erde hat ein relativ starkes Magnetfeld, das in 1. Näherung durch einen magnetischen Dipol mit dem magnetischen Moment von $8{,}06 \cdot 10^{15}\,\mathrm{Tm}^{-3}$ beschrieben werden kann. Der Dipol muß zur Erklärung des Magnetfelds exzentrisch angeordnet werden. Seine Achse ist gegenüber der Rotationsachse um 11,5° geneigt und durchstößt die Nordhalbkugel bei 78,5° N, 291° O und die Südhalbkugel bei 78,5° S, 111° O. Der magnetische Nordpol liegt auf der Südhalbkugel, d. h. im Feldlinienbild treten die magnetischen Kraftlinien auf der Südhalbkugel aus und laufen zur Nordhalbkugel. Das Erdmagnetfeld wird an der Erdoberfläche durch ein umfangreiches Observatoriumsnetz erfaßt und enthält Anteile verschiedener Quellen. Der größte Anteil, das erdmagnetische Hauptfeld, hat seine Quellen im Erdinneren und zeigt eine sehr geringe zeitliche Veränderlichkeit, die *Säkularvariation*. Ein sehr viel kleinerer, zeitlich stark variabler Teil, das *Variationsfeld*, hat seine Quellen in der Ionosphäre (↑ 10.1.3.) und in der Magnetosphäre (↑ 10.1.4.).

Das *erdmagnetische Hauptfeld* zeigt großräumige Änderungen über Tausende von Kilometern, die nicht mit geologischen Strukturen korreliert sind, ihren Ursprung also in Tiefen haben müssen, die mit geologischen Methoden nicht mehr erfaßt werden. Aus den vorliegenden Meßreihen wurden zeitliche Variationen mit Perioden von einigen hundert bis einigen tausend Jahren erschlossen. Untersuchungen von Gesteinen, deren Bausteine sich magnetisch ausrichten und die die bei ihrer Entstehung herrschenden Magnetverhältnisse konservieren, ergaben, daß das Hauptfeld in der Vergangenheit in erdgeschichtlich relativ kurzen Zeitspannen seine Polarität mehrfach änderte und dazwischen relativ konstant blieb. Das ist nur erklärbar, wenn man annimmt, daß die Quellen des Hauptfelds im flüssigen äußeren Erdkern liegen.

Treten im mit hoher elektrischer Leitfähigkeit (10^{-4} bis $10^{-5}\,\Omega^{-1}\mathrm{m}^{-1}$) versehenen flüssigen äußeren Erdkern Strömungen mit einer Geschwindigkeit von etwa $10^{-3}\,\mathrm{cm\,s^{-1}}$ auf, so kann die Existenz des Hauptfelds durch die *Dynamotheorie* plausibel gemacht werden: Es existiere ein primäres Magnetfeld. Bewegt sich in diesem Magnetfeld eine elektrisch leitende Flüssigkeit, so kann es bei geeigneter Konfiguration und Strömung zu einer Verstärkung des Magnetfelds, d. h. zu einer Umwandlung von mechanischer Strömungsenergie in magnetische Energie kommen. Auch für das weitere Aufrechterhalten des Magnetfelds ist eine entsprechende Strömung erforderlich. Das Problem des magnetohydrodynamischen Dynamos ist wegen der starken Verkopplung der einzelnen Größen sehr kompliziert. Das Strömungsfeld entsteht unter der Wirkung thermodynamischer und physikochemischer Prozesse, aber auch elektrodynamischer Kräfte. Ein Anwachsen des Magnetfelds führt daher nach einiger Zeit zu einem Anwachsen dieser Kräfte und einer Rückwirkung auf die Strömung. Unter starker Wirkung nichtlinearer Prozesse müßte sich nach und nach ein Gleichgewicht mit wenig veränderlichem Magnetfeld einstellen. Bisher beschränkte man sich auf die elektrodynamische Seite des Problems und untersuchte die Eigenschaften der entstehenden Magnetfelder in Abhängigkeit von vorgegebenen Strömungen. Dabei fand man, daß eine axialsymmetrische Strömung kein axialsymmetrisches Magnetfeld aufrechterhalten kann und daß sowohl das Magnetfeld als

auch das Strömungsfeld einen gewissen Grad an Komplexität haben müssen, damit der magnetohydrodynamische Generator (↑ 9.1.8.5.) läuft. Andererseits zeigt der Erdkern aber gerade einen hohen Grad von Symmetrie. STEENBECK, KRAUSE und RÄDLER haben in den letzten Jahren gezeigt, daß turbulente Strömungen niederer Ordnung geeignete Magnetfelder erzeugen können. Tritt z. B. unter der Wirkung der Corioliskraft eine elektromotorische Kraft parallel zur Richtung des mittleren Magnetfelds auf, so wird ein erdfeldähnliches Feld erzeugt. Eine ähnliche Wirkung hat die Existenz einer differentialen Rotation, d. h. das Auftreten einer Tiefenabhängigkeit der Winkelgeschwindigkeit. Von einer quantitativen Deutung der Entstehung und Änderung des geomagnetischen Hauptfelds ist man jedoch noch weit entfernt.

10.1.2. Physik der Neutralgasatmosphäre

Am Erdboden besteht die Atmosphäre der Erde zu 75,52% aus N_2, zu 23,14% aus O_2, zu 1,29% aus Ar und zu 0,05% aus CO_2. Sie hat insgesamt eine Masse von $5{,}13 \cdot 10^{17}$ kg. Unter der Wirkung der Schwerkraft, der eingestrahlten Sonnenenergie und zahlreicher dynamischer Prozesse bildet sich eine komplizierte Struktur aus. Bis zu einer Höhe von 105 km reicht die Homosphäre mit konstanter mittlerer rel. Molmasse $\bar{M}$. Darüber in der Heterosphäre entmischen sich die Gase und $\bar{M}$ nimmt ab (Abb. 10.1.2-1), die leichteren Gasbestandteile überwiegen. Zugleich nimmt die mittlere freie Weglänge $\bar{l}$ zu. In der *Heterosphäre* führt die Fotodissoziation molekularen Sauerstoffs durch ultraviolettes Licht mit $\lambda < 220$ nm zu atomarem Sauerstoff. Fotochemische Prozesse können auch die Dichte von molekularem Stickstoff vermindern. Die Beobachtung der Temperatur T bildet die Grundlage für eine feinere Gliederung. Jede der dabei definierten Sphären wird von einer als *Pause* bezeichneten Grenzschicht nach oben hin abgeschlossen. So ist z. B. die *Tropopause* (12 bis 15 km) die obere Grenze der *Troposphäre*. Die Temperatur nimmt in der Troposphäre zunächst mit der Höhe ab und erreicht in der Tropopause ein Minimum. In der darauffolgenden *Stratosphäre* nimmt sie zu und erreicht in der *Stratopause* (etwa 50 km) ein Maximum von etwa 270 K; danach in der *Mesosphäre* sinkt sie stark ab und erreicht in der *Mesopause* (85 bis 90 km) das absolute Temperaturminimum der Erdatmosphäre von etwa 180 K. Oberhalb 90 km in der *Thermosphäre* steigt die Temperatur zunächst stark an und zeigt dann oberhalb von 300 bis 400 km keine Höhenabhängigkeit mehr. Dieser höhenunabhängige Temperaturwert wird als *Exosphärentemperatur* bezeichnet (↑ Abb. 10.1.2-1).

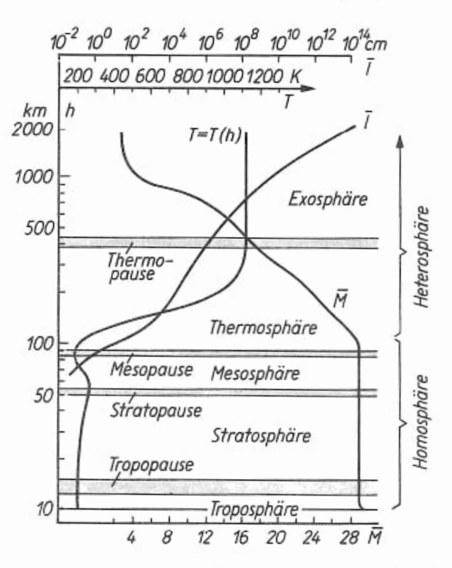


Abb. 10.1.2-1 Neutralgasparameter während ruhiger Bedingungen; Temperatur T, mittleres Molekulargewicht $\bar{M}$, mittlere freie Weglänge $\bar{l}$ als Funktion der Höhe h

10.1.2.1. Untere Atmosphäre. Die physikalischen Verhältnisse in der unteren Atmosphäre werden dadurch gekennzeichnet, daß der primäre Umsatz infraroter und sichtbarer Strahlung in Wärme an der Erdoberfläche und an Strukturen der untersten Schichten erfolgt und anschließend vorwiegend thermodynamische Transportprozesse auftreten. Hier findet das Wettergeschehen statt, für das die Bewegung von Luftmassen als Folge der Absorption der Strahlungsenergie der Sonne von entscheidender Bedeutung ist. Dabei müssen Absorptions-, Reflektions- und Streuprozesse am Erdboden, an den Wolken und an Aerosolen sowie der Transport der Energie durch Leitung, Konvektion und Kondensation berücksichtigt werden. Trotz umfangreicher Messungen ist die Energiebilanz dieses Gebiets noch nicht genau bekannt.

10.1.2.2. Hochatmosphäre. In der Hochatmosphäre wird das gesamte ultraviolette Spektrum der Sonne absorbiert. Die absorbierte Energie wird durch fotochemische Prozesse wie Dissoziation und Ionisation, die selbst noch zusätzliche Energie abgeben, oder durch Folgeprozesse wie Rekombination überwiegend als Wärme freigesetzt. Die Energieaufnahme der Hochatmosphäre wird von allen Variationen der Dichte und Zusammensetzung der absorbierenden Gaskomponenten beeinflußt. Das Temperaturverhalten der Hochatmosphäre kann nur unter Hinzuziehung von dynamischen Pro-

zessen, die für eine horizontale und vertikale Umverteilung der Energie sorgen, verstanden werden.

10.1.2.2.1. *Struktur der Hochatmosphäre.* Das umfangreiche Datenmaterial ist in Standardatmosphären zusammengefaßt worden, z. B. in

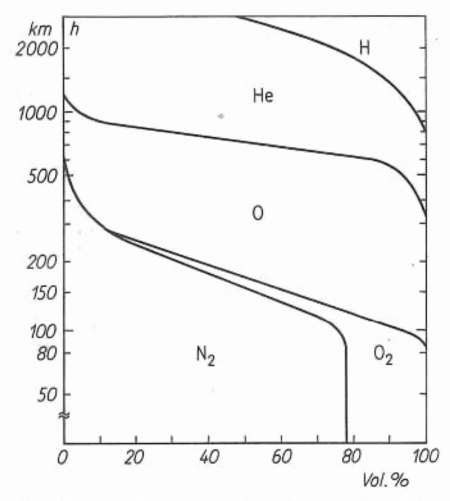


Abb. 10.1.2-2 Prozentuale mittlere chemische Zusammensetzung der Atmosphäre als Funktion der Höhe *h* nach CIRA 72

dem CIRA-Modell [*C*ommittee on Space Research, *I*nternational *R*eference *A*tmosphere, engl., Komitee für Weltraumforschung, Internationale Bezugsatmosphäre]. Die chemische Zusammensetzung der Atmosphäre in Abhängigkeit von der Höhe nach der CIRA 1972 zeigt Abb. 10.1.2-2. Dabei ist die Zusammensetzung in Volumenprozent angegeben, die Linien trennen verschiedene chemische Elemente. Man erkennt, daß oberhalb 90 km Höhe die molekularen Bestandteile stark gegenüber dem atomaren Sauerstoff zurücktreten. In ungefähr 200 km beträgt der Anteil des atomaren Sauerstoffs bereits etwa 60%, der des molekularen Stickstoffs nur noch 35% und der des molekularen Sauerstoffs 5%. In Höhen oberhalb 500 km tritt der Einfluß des atomaren Sauerstoffs im Vergleich zum Helium zurück, das dann in noch größeren Höhen durch atomaren Wasserstoff als Hauptbestandteil abgelöst wird. In ihrer allgemeinen Struktur entspricht diese chemische Zusammensetzung der Thermosphäre dem aus der diffusiven Entmischung und Fotodissoziation zu erwartenden Verhalten. Daneben findet man jedoch starke jahreszeitliche Änderungen, z. B. im Konzentrationsverhältnis des atomaren zum molekularen Sauerstoff oder in der Dichteverteilung des Heliums, die die Folge komplexer jahreszeitlicher Zirkulationsvorgänge in der Hochatmosphäre sein dürften.

Die *Temperaturverteilung in der Strato- und Mesosphäre* ist für Sonnenwendbedingungen als Meridianschnitt in Abb. 10.1.2-3 in Abhängigkeit von der Höhe und Breite durch Kurven gleicher Temperatur angegeben. Im Sommer ändert sich die Temperatur mit zunehmender Höhe und Breite sehr viel stärker als im Winter.

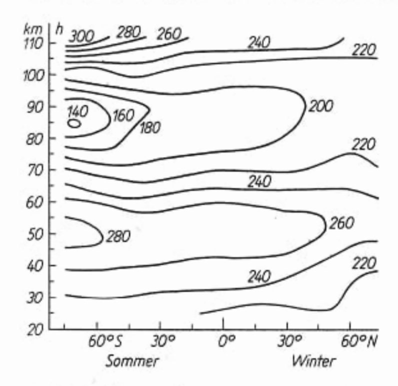


Abb. 10.1.2-3 Isothermen der Strato- und Mesosphäre während des Solstitiums nach CIRA 72; Temperaturangaben in Kelvin

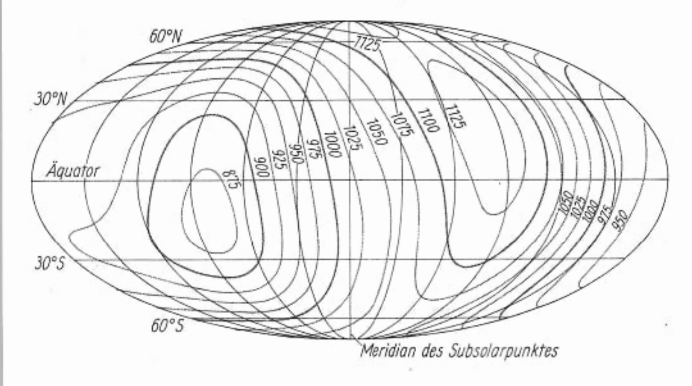


Abb. 10.1.2-4 Globale Verteilung der Exosphärentemperatur aufgrund direkter Messungen; Linien gleicher Temperatur in K

In der Mesopause ist sie im Sommer beträchtlich niedriger als im Winter, in 50 km Höhe ist sie dagegen im Sommer höher als im Winter. Mit zunehmender Höhe verschiebt sich die Lage des Punktes tiefster Temperatur vom Dezember in 50 km zum Juni in 90 km Höhe; die Amplitude dieser jahreszeitlichen Variation ist wesentlich größer als die möglicher Tagesgänge. Tageszeitliche Variationen der Temperatur dürften in diesem Höhengebiet 5% nicht überschreiten.

Die *Temperaturverteilung in der Thermosphäre* konnte erst vor kurzem durch neue Meßverfahren direkt erfaßt werden. Abb. 10.1.2-4 zeigt die globale Verteilung der Exosphärentemperatur. Die Sonneneinstrahlung ist am Subsolarpunkt, d. h. dem Punkt der Erdoberfläche, über dem die Sonne genau im Zenit steht, am stärksten. Das Maximum der Exosphärentemperatur liegt zwar auf der gleichen geographischen Breite wie der Subsolarpunkt, ist jedoch um etwa 70 Längengrade verschoben, d. h. es tritt erst bei 16 bis 17 Uhr Ortszeit auf. Die Exosphärentemperatur nimmt nach allen Richtungen hin mit wachsender Entfernung vom Subsolarpunkt ab.

Globale Verteilungen der Exosphärentemperatur wurden im großen Umfang auch aus experimentell bestimmten globalen Verteilungen der Dichte hergeleitet. In diesen statischen Modellen wird die Temperaturverteilung berechnet, die unter der Voraussetzung diffusiven Gleichgewichts die Dichtemessungen am besten erklären kann. In der CIRA 72 wurde das auf der Grundlage von Gesamtdichtemessungen durchgeführt. Dabei findet man das Maximum der Exosphärentemperatur zwar auch auf der Subsolarpunktbreite, es läuft dem Subsolarpunkt jedoch nur um 30 Längengrade nach. Die tiefste Exosphärentemperatur liegt danach auf der Nachtseite in der Nähe des Antipodenpunkts des Temperaturmaximums. Die Temperaturverteilung zeigt also einen ausgesprochenen tageszeitlichen Buckel [engl. *diurnal bulge*]. Dreht sich ein Beobachter mit der Erde unter dieser Temperaturverteilung durch, so stellt er einen Tagesgang der Exosphärentemperatur fest, dessen Amplitude und Mittelwert um so kleiner werden, je weiter die geographische Breite des Beobachtungsorts von der des subsolaren Punkts entfernt ist. Diese statischen Temperaturmodelle müssen jedoch stets als Umrechnungen beobachteter Dichteverteilungen verstanden werden.

Die Tatsache, daß das beobachtete Temperaturmaximum nicht mit dem beobachteten Dichtemaximum zusammenfällt, zeigt, daß die Temperatur- und Dichteverteilungen nicht statisch, sondern nur bei Beachtung interner Winde und Strömungen, also dynamischer Prozesse, verstanden werden können.

Die in Abb. 10.1.2-4 gezeigte Temperaturverteilung baut ein Druckgefälle auf, das die Luft von der Tagesseite auf die Nachtseite abfließen läßt. Dieser Abtransport kann den Dichteanstieg aufgrund des Temperaturanstiegs überwiegen. Dadurch nimmt trotz weiter ansteigender Temperatur die Dichte ab, und das Dichtemaximum tritt zu einem früheren Zeitpunkt auf als das Temperaturmaximum.

Die Temperatur hängt ebenso wie andere thermosphärische Größen von der *Energieeinstrahlung*, d. h. von der Sonnenaktivität und der magnetischen Aktivität ab. Sie folgen sowohl langfristigen als auch kurzfristigen Variationen der Sonnenaktivität. Während die maximale Exosphärentemperatur im Sonnenfleckenminimum bei 800 K liegen kann, steigt sie im Maximum der Sonnenaktivität auf 1300 bis 1500 K an.

Während Energieerhöhungen durch Einstrahlung solarer elektromagnetischer Strahlung in allen Breiten gleichzeitig wirksam werden, werden die durch verstärkten Sonnenwind in die Magnetosphäre übertragenen Energien (s. a. 10.1.3.3.4.) nur in hohe Breiten eingespeist. Dadurch entsteht ein Nachhinken der thermosphärischen Erwärmung im Vergleich zum Anstieg der geomagnetischen Aktivität, das in höheren Breiten etwa 1,5 h, in niedrigeren geographischen Breiten dagegen 4,5 bis 6 h beträgt. Die Erwärmung breitet sich also aus der Atmosphäre der Polarlichtzone durch Transportprozesse zu niedrigen Breiten hin aus. Welche thermodynamischen Transportvorgänge hier wirksam werden und so die Kopplung der Erdatmosphäre mit extraterrestrischen Einwirkungen bedingen, ist vorläufig unklar.

Der Einfluß der Sonnenaktivität auf die Struktur und Dynamik der Thermosphäre ist eindeutig nachgewiesen, unklar ist jedoch, wie weit der vertikale Durchgriff der Sonnenaktivitätsvariationen in die Erdatmosphäre hineinreicht.

10.1.2.2.2. *Strahlungsbilanz der Hochatmosphäre.* Das gesamte ultraviolette Sonnenspektrum mit $\lambda < 290$ nm wird in der Hochatmosphäre absorbiert. Wegen der sehr geringen Masse der Hochatmosphäre kontrolliert dieser in bezug auf die Gesamtenergie der in die Atmosphäre einfallenden Wellenstrahlung geringe Energiebetrag vollständig die charakteristischen physikalischen Eigenschaften der Hochatmosphäre.

Die *spektrale Verteilung der Sonnenstrahlung* im ultravioletten Bereich zeigt typische stark über das kontinuierliche Spektrum hervortretende Spektrallinien. Die Intensität dieser Emissionslinien schwankt stark in Abhängigkeit von der Sonnenaktivität. Die Absorption dieses Spektrums in der Atmosphäre erfolgt in der Hauptsache durch molekularen Sauerstoff und durch Ozon. Erst bei Wellenlängen $\lambda < 100$ nm

kommt in merklichem Umfang Absorption durch atomaren Sauerstoff und durch Stickstoff hinzu. Strahlung dieser Wellenlängen führt unmittelbar zur Ionisation der Neutralgasteilchen und zur Bildung der Ionosphäre (↑ 10.1.3.1.).
Im Bereich 195 bis 320 nm wird die Strahlung durch Ozon absorbiert. Da der Strahlungsfluß in diesem Bereich stark ansteigt, stellt trotz der geringen Menge des Ozons dieser Absorptionsprozeß die *Hauptwärmequelle für die obere Stratosphäre und Mesosphäre* dar. Das Maximum des atmosphärischen Ozons findet man in einer Höhe von etwa 25 km mit $5 \cdot 10^{18}$ Teilchen pro m^3. Aus den vertikalen Verteilungen von O_3, O_2 und O kann die pro Volumeneinheit und Sekunde absorbierte Sonnenstrahlungsenergie als Funktion der Höhe berechnet werden. Das Maximum der gesamten Strahlungsabsorption liegt in etwa 40 km Höhe, also oberhalb des Konzentrationsmaximums des O_3 und unterhalb des Temperaturmaximums der Stratopause.
Die *Energieaufnahme der Hochatmosphäre* wird vor allem durch zeitliche und räumliche Veränderungen der Dichte und Höhenverteilung der absorbierenden Gaskomponenten beeinflußt. Um die *Energiebilanz* aufstellen zu können, muß man auch die Wärmesenken des Strahlungsenergieumsatzes kennen. In der oberen Stratosphäre und der Mesosphäre stellt die infrarote Strahlung des Kohlendioxidmoleküls bei 15 μm den hauptsächlichsten *Energieverlustprozeß* dar. Kleinere Beiträge zur infraroten Energieabstrahlung liefern in der Stratosphäre auch noch die O_3-, H_2- und O_2-Moleküle. In der Thermosphäre kommt als zusätzlicher Abstrahlungsprozeß die infrarote 63-μm-Linie des atomaren Sauerstoffs hinzu. Berechnet man die Erwärmungs- bzw. Abkühlungsraten, so stellt man fest, daß das Temperaturverhalten der Hochatmosphäre nicht durch Strahlungsgleichgewicht erklärt werden kann. Vielmehr spricht alles dafür, daß Zirkulations- und andere dynamische Prozesse für eine starke Umverteilung der Energie sorgen (Tafeln 52/53).
10.1.2.2.3. *Dynamik der Hochatmosphäre.* In allen Stockwerken der Hochatmosphäre spielen *Zirkulationssysteme* eine besondere Rolle. Das *strato-mesosphärische Zirkulationssystem* wird ähnlich wie die Zirkulation in der unteren Atmosphäre durch die jahreszeitlich-geographischen Temperaturgradienten angetrieben. Es besteht in einer vorherrschenden westöstlichen Strömung rund um die Pole herum. Diese zonalen in Richtung nach Westen laufenden gleichmäßigen Winde sind während der Tagundnachtgleichen und im Sommer ausgeprägt. Im Winter treten beträchtlich stärkere, stark veränderliche Winde in Richtung nach Osten auf. Das strato-mesosphärische Zirkulationssystem ist mit steigender Höhe zunehmend mit *Gezeitenschwingungen* überlagert. Es schließt nach oben ziemlich genau mit der Mesopause ab.
Das *Zirkulationssystem der Thermosphäre* wird durch die tagszeitlichen Druck- bzw. Temperaturvariationen angetrieben. Aufgrund der hohen kinematischen Zähigkeit und des Auftretens der *Ionenreibung* reagiert das thermosphärische Neutralgas jedoch anders als das der Mesosphäre. Unter Ionenreibung versteht man die abbremsende Wirkung der Ionen auf das Neutralgas, die dadurch hervorgerufen wird, daß die Ionen schon mehr oder weniger fest an die magnetischen Kraftlinien gebunden sind und der Neutralgasbewegung nicht mehr frei folgen (↑ 10.1.3.1.2.). Das Zirkulationssystem der Thermosphäre zeigt eine tageszeitlich um die Erde laufende Strömung von der Tag- auf die Nachtseite, die auch über die Pole hinweg läuft, sowie eine resultierende mittlere Strömung in Richtung von Westen nach Osten, die oft als *Superrotation der Thermosphäre* bezeichnet wird. Wegen der hohen inneren Reibung fließt die Strömung in Richtung der Druckgradienten.
Die Mesopausenregion und die untere Thermosphäre stellen das *Übergangsgebiet zwischen den beiden Zirkulationssystemen dar.* Für die meiste Zeit des Jahres scheint die Mesopause an das Zirkulationssystem der Thermosphäre angeschlossen zu sein, nur während gewisser Übergangszeiten erfolgt ein Hereingreifen des mesosphärischen Zirkulationssystems in die Mesopausenregion. Gezeitenwinde spielen in der Mesopausenregion und in der unteren Thermosphäre eine wesentliche Rolle.

10.1.3. Physik der Ionosphäre

Die Energie des kurzwelligen Teils des Sonnenspektrums (↑ 10.1.2.2.2.) und der hochenergetischen Partikel kosmischen, solaren oder magnetosphärischen Ursprungs reicht aus, um die Atome und Moleküle des Neutralgases der Meso- und Thermosphäre zu ionisieren. Das entstehende schwach ionisierte Plasma wird als *Ionosphärenplasma* bezeichnet. Die *Ionosphäre* stellt ein Übergangsgebiet zwischen den neutralen unteren Atmosphärenschichten und dem vollionisierten Plasma der Magnetosphäre dar.
Im untersten Teil der Ionosphäre spielen *Stöße zwischen Partikeln* die dominierende, im oberen Teil eine kaum noch zu beachtende Rolle. Im oberen Teil bestimmen *elektromagnetische Kräfte*, vor allem das Magnetfeld durch die Lorentzkraft, das physikalische Geschehen, deren Wirkung mit abnehmender Höhe, d. h. zunehmender Dichte, an Bedeutung verliert.
Neben dieser physikalischen, die Abgrenzung zur Magnetosphäre betonenden Definition, gibt es die historisch gewachsene Definition. Da-

nach werden als *Ionosphäre* die Schichten ionisierter Atmosphäre bezeichnet, die vom Boden aus mit Hilfe elektromagnetischer Wellen sondiert werden können (60 bis etwa 400 km).
Dringt von außen energiereiche Strahlung in die nach unten immer dichter werdende Erdatmosphäre ein, so erzeugt sie eine um so höhere Dichte freier Elektronen, je tiefer sie kommt. Da beim Ionisationsprozeß aber Energie absorbiert wird, muß es ein *Niveau maximaler Elektronenproduktion* geben, unterhalb dessen die Erzeugungsrate der Elektronen trotz weiteren Dichteanstiegs der Neutralteilchen abnehmen wird. Die Höhe maximaler Elektronenproduktion ist unabhängig von der Intensität der einfallenden Strahlung. Bis in welche Höhe hinab jedoch die ionisierende Strahlung vordringen kann, hängt von der Dichte und Zusammensetzung des Neutralgases ab. Da die Prozesse, die zur Rekombination, d. h. zum Verschwinden freier Ladungsträger, führen, in noch stärkerem Maße vom Neutralgas abhängen, verhält sich die Ionosphäre im Thermosphärenbereich anders als im Mesosphärenbereich.

10.1.3.1. Ionisation der Thermosphäre. Für die Ionisation von atomarem Sauerstoff benötigt man Licht mit Wellenlängen $\lambda \leqq 91$ nm, für O_2-Moleküle $\lambda = 102{,}7$ nm und für N_2-Moleküle mit $\lambda = 79{,}6$ nm. Abb. 10.1.3-1 zeigt die Elektronenerzeugungsrate in Abhängigkeit von der Höhe. Man erkennt, daß praktisch die ganze Strahlung mit $\lambda \leqq 102{,}7$ nm (mit Ausnahme der harten Röntgenstrahlung $\lambda \leqq 0{,}5$ nm) in der Thermosphäre verbraucht wird.
Das Verhalten der *thermosphärischen Ionisation* wird seit mehr als 40 Jahren mit Hilfe der elektromagnetischen Impulsecholotung weltweit untersucht. Es zeigen sich zwei Elektronendichtemaxima in 100 bis 120 km Höhe und in 250 bis 350 km Höhe, welche die jeweils obere Grenze zweier Hauptgebiete der thermosphärischen Ionisation, der *E-Region* (früher Heavisideschicht) und der *F-Region* (früher Appletonschicht), darstellen (Abb. 10.1.3-2).

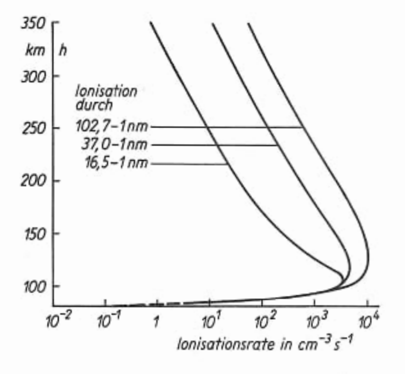


Abb. 10.1.3-1 Elektronenproduktionsraten unter ungestörten Bedingungen; *h* Höhe

E- und F-Region zeigen sowohl bezüglich der Variationen mit der Zeit als auch mit der Breite ein völlig unterschiedliches Verhalten. Die Elektronendichtewerte der E-Region lassen sich in

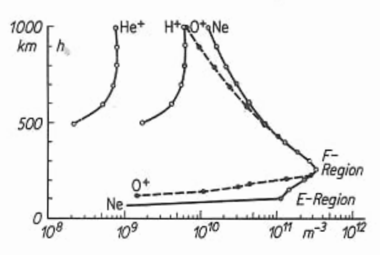


Abb. 10.1.3-2 Elektronendichte und Ionenzusammensetzung für die Mittagsstunde über einem Ort 45° nördlicher geographischer Breite nach IRI 77

guter Näherung als Funktion des Zenitwinkels der Sonne, also der Energieeinstrahlung verstehen. So wird z. B. eine tägliche Variation der Elektronendichte mit einem Maximum um die Mittagszeit beobachtet. Massenspektrometermessungen zeigen, daß in der E-Region molekulare Ionen wie O_2^+ und NO^+ überwiegen (↑ Abb. 10.1.3-2), wobei letztere durch Ionenreaktionen aus den primären N_2^+-Ionen entstanden sind. Diese molekularen Ionen rekombinieren so schnell mit den freien Elektronen, daß sich ein ständiges Quasigleichgewicht zwischen Elektronenerzeugungs- und -verlustprozessen einstellt und keine Zeit zum Transport der freien Ladungen verbleibt. In der F-Region überwiegen atomare Ionen wie O^+ (↑ Abb. 10.1.3-2). Die direkte Rekombination atomarer Ionen mit freien Elektronen läuft sehr langsam ab. Effektiver sind Prozesse, bei denen die atomaren Ionen ihre positive Ladung an Moleküle abgeben wie $O^+ + O_2 \rightarrow O + O_2^+$ oder $O^+ + N_2 \rightarrow N + NO^+$. Diese molekularen Ionen rekombinieren dann schneller durch dissoziative Rekombination mit freien Elektronen. Die Geschwindigkeit der Verlustprozesse hängt somit von der Dichte der molekularen Bestandteile des Neutralgases ab. Da die Dichte mit zunehmender Höhe ständig abnimmt, erfolgt der Verlust freier Elektronen immer langsamer und das Maximum der Elektronendichte fällt nicht mit dem Maximum der Elektronenproduktion zusammen, sondern liegt wesentlich höher (↑ Abb. 10.1.3-2). So spiegeln Aufbau und Verhalten der thermosphärischen Ionisationsgebiete die chemische Zusammensetzung der Thermosphäre und deren Variationen wider.
Aufgrund der langen Lebensdauer der Elektronen in der F-Region spielen bei der Einstellung

des Gleichgewichts nicht nur Erzeugungs- und Verlustprozesse, sondern auch verschiedene *dynamische Prozesse* eine entscheidende Rolle. In Frage kommen im wesentlichen drei Prozesse: Diffusion, Transport von Ladungsträgern unter der Wirkung thermosphärischer Neutralgaswinde und Transport unter der Wirkung dynamo-elektrischer (↑ 10.1.3.2.) oder magnetosphärischer elektrischer Felder. Mit zunehmender Höhe wird das erdmagnetische Feld mehr und mehr zur bewegungsbestimmenden Kraft für die elektrisch geladenen Teilchen. Daher rufen Bewegungen des Neutralgases (Winde), die im wesentlichen horizontal verlaufen, kein einfaches horizontales Mitnehmen des Ionen- bzw. Elektronengases hervor. Vielmehr löst wegen der Neigung der magnetischen Kraftlinien ein horizontaler Stoß eines Neutralteilchens auf ein Ion durch die Wirkung der Lorentzkraft eine Gyrationsbewegung um die Feldlinien und je nach dem Richtungssinn der magnetfeldparallelen Geschwindigkeitskomponente eine Aufwärts- oder Abwärtsbewegung des geladenen Teilchens längs der Magnetfeldlinie aus. Ein Neutralgaswind bewirkt also in erster Linie eine vertikale Verlagerung der Ladungsträger. Praktisch die gleiche Wirkung erzielt auch ein horizontales elektrisches Feld.

Die den Teilchen erteilte Geschwindigkeit führt also zu einem Transport längs der magnetischen Kraftlinie. So erzeugt auf der Nordhalbkugel ein nordwärts wehender Neutralgaswind, wie er tagsüber existiert, eine Abwärtsbewegung der Ionen und Elektronen. Diese kommen damit in ein Gebiet erhöhter effektiver Rekombination, die Ionendichte nimmt ab. Weht dagegen ein südwärts gerichteter Wind, so werden Elektronen und Ionen in größere Höhen, d. h. in Gebiete verlangsamter Rekombinationsprozesse, transportiert.

Die thermosphärische Ionisationsverteilung spiegelt nicht nur den Aufbau und die Zusammensetzung des Neutralgases, sondern auch seine Bewegungsvorgänge in komplizierter Art und Weise wider. Unter Berücksichtigung dieser Vorgänge wird es möglich, die komplizierten tageszeitlichen und breitenabhängigen Variationen der Elektronendichte der F-Region qualitativ verständlich zu machen.

10.1.3.2. Elektrische Felder und Leitfähigkeiten in der Ionosphäre. In der Ionosphäre oberhalb 90 km wird die Bewegung der Ionen und Elektronen sowohl durch das erdmagnetische Feld als auch durch Stöße mit den Neutralteilchen bestimmt. Da die Elektronen eher der magnetischen Feldwirkung folgen, d. h. fester durch das Magnetfeld geführt werden als die Ionen, und die Stöße zwischen Ionen und Neutralteilchen wirksamer sind, wirken auf Ionen und Elektronen verschiedene Kräfte; es kann zur *Ladungstrennung* kommen. Das entstehende *dynamo-elektrische Feld* hängt von dem den Prozeß treibenden Neutralgaswind und der *stark anisotropen elektrischen Leitfähigkeit* ab (Dynamomechanismus).

Bei der Betrachtung der Komponenten der elektrischen Leitfähigkeit hat man zwischen der *Spitzerleitfähigkeit* σ_0 parallel zum Magnetfeld, der *Pedersenleitfähigkeit* σ_1 senkrecht zum Magnetfeld und der *Halleitfähigkeit* σ_2 senkrecht zum elektrischen und zum Magnetfeld zu unterscheiden. Abb. 10.1.3-3 zeigt die Höhenabhängigkeit der Komponenten des *Leitfähigkeitstensors*. Die parallele Leitfähigkeit nimmt mit zunehmender Höhe schnell zu und erreicht so hohe Werte, daß die Annahme $\sigma_0 \rightarrow \infty$ oft berechtigt ist. Für großräumige elektrische Felder und Ströme existiert deshalb ständig ein Quasikurzschluß, d. h., die Feldlinien können im allgemeinen als Äquipotentiallinien betrachtet werden. Im Gebiet von 100 bis 150 km Höhe erreichen die Komponenten der Leitfähigkeit senkrecht zum Magnetfeld vergleichbare Werte. Nur in diesem Gebiet der Magnetosphäre kann ein nennenswerter Strom senkrecht zum erdmagnetischen Hauptfeld fließen.

In Höhen um 100 und 140 km zeigt die Verteilung der senkrechten Komponenten der elektrischen Leitfähigkeit starke Gradienten, die für die Erzeugung großräumiger ionosphärischer elektrischer Felder eine Rolle spielen.

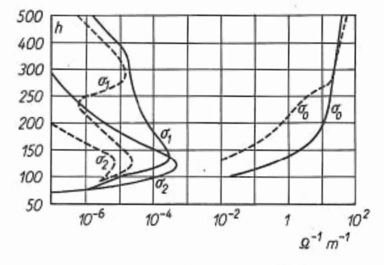


Abb. 10.1.3-3 Spitzer- (σ_0-), Pedersen- (σ_1-) und Hall- (σ_2-) Leitfähigkeiten in Abhängigkeit von der Höhe h in km für Ortsmittag und Ortsmitternacht (*gestrichelt*)

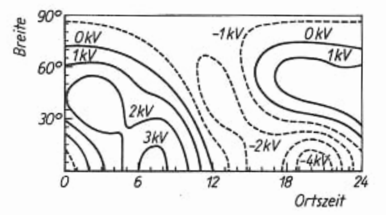


Abb. 10.1.3-4 Verteilung der elektrischen Feldstärke bei ungestörten Bedingungen nach Richmond

Abb. 10.1.3-4 zeigt die *globale Verteilung der dynamo-elektrischen Felder* nach RICHMOND. Diese elektrischen Felder führen sowohl zur vertikalen Verlagerung von Ladungsträgern als auch zum Aufbau großräumiger die ganze Erde umfassender *Stromsysteme*, die von der Sonne aus gesehen ortsfest sind. Die magnetische Wirkung dieses Stromsystems spiegelt sich für ein sich mit der Erde drehendes Observatorium als ein tageszeitlich veränderliches Magnetfeld wider. Diese an ruhigen Tagen gut zu beobachtenden S_q-Variationen sind ein Teil des erdmagnetischen Variationsfelds.

10.1.3.3. Ionisation der Mesosphäre. Da die energiereiche Sonnenstrahlung mit $\lambda \leqq 102{,}7$ nm mit Ausnahme der kurzwelligen Röntgenstrahlung in Höhen oberhalb 90 km vollständig absorbiert wird (↑ 10.1.3.1.) und die Röntgenstrahlung nur bei starken Störungen auf der Sonne eine beachtliche Intensität erreicht, kann die Ionisation der normalen Luftbestandteile unterhalb 90 km keine wesentliche Rolle mehr spielen. Für die Erzeugung freier Elektronen und Ionen sowie für die ablaufenden Reaktionen spielen Spurenstoffe wie NO, CO_2 oder H_2O eine entscheidende Rolle. Man nimmt heute an, daß die freien Elektronen unterhalb der Mesopause vorwiegend durch Fotoionisation von Stickoxid NO, ionisierbar durch $\lambda \leqq 134$ nm, erzeugt werden.

Die beobachteten Teilchendichten von einigen 10^7 NO-Molekülen pro cm^3 und die Energie und Intensität der solaren Lyman-α-Linie (121,6 nm) reichen aus. Von den zur Bildung von NO-Molekülen führenden Prozessen dürfte die Reaktion von angeregten Stickstoffatomen $N(^2D)$ mit O_2-Molekülen am effektivsten sein. Die angeregten $N(^2D)$-Atome könnten in den thermosphärischen Ionisationsgebieten gebildet und durch Diffusion in tiefere Schichten gelangt sein. Angeregte metastabile Atome und Moleküle spielen bei zahlreichen chemischen und ionenchemischen Reaktionen in der Mesosphäre und Thermosphäre eine bedeutende Rolle.

In den mesosphärischen Ionisationsgebieten sind auch andere, sonst kaum beobachtete Partikel von Bedeutung. Massenspektrometrische Messungen haben gezeigt, daß unterhalb 86 bis 82 km die sonst üblichen O_2^+- und NO^+-Ionen plötzlich verschwinden und statt dessen sogenannte *Wasser-Cluster-Ionen* als positive Ladungsträger auftreten, z. B. $NO^+(H_2O)_n$, $O_2^+ \cdot (H_2O)_n$ und $H^+ \cdot (H_2O)_n$. Die Rekombination der freien Elektronen mit diesen Wasser-Cluster-Ionen erfolgt effektiver.

Im Höhenbereich unterhalb 80 km bestehen die

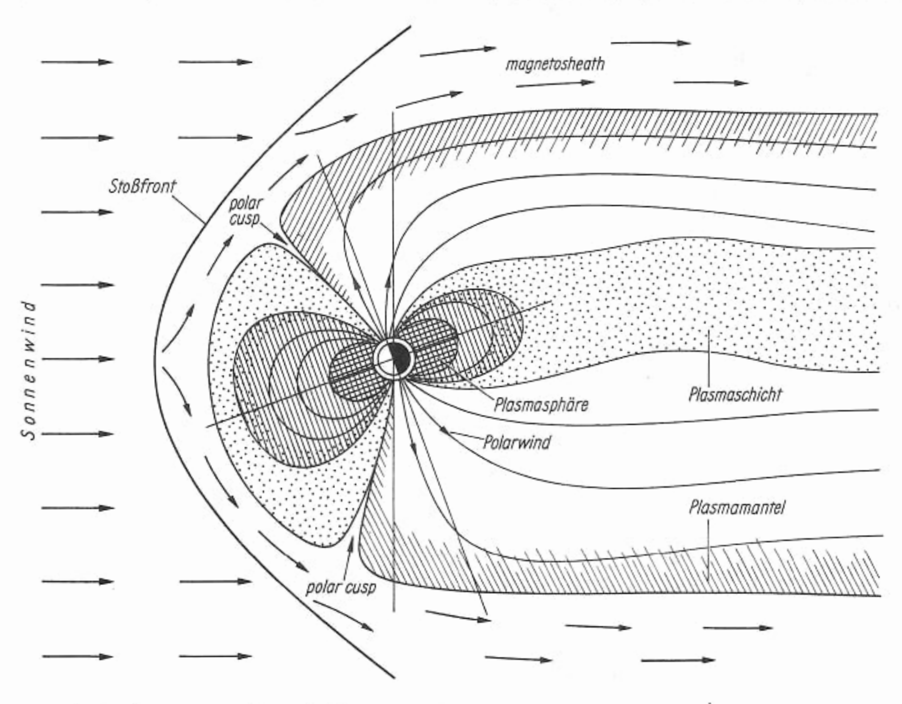


Abb. 10.1.4-1 Magnetfeld und Hintergrundplasma in der Magnetosphäre („polar cusp" polare Hörner")

negativen Ladungsträger nachts nicht mehr überwiegend aus freien Elektronen, sondern aus negativen Molekülionen O_2^-, CO_3^-, NO_3^-. Diese entstehen, wenn durch Anlagerung eines freien Elektrons an ein Neutralteilchen ein quantenmechanisch stabiles System entsteht. Allerdings reicht zur Abspaltung des Elektrons schon die Energie des sichtbaren Sonnenlichts aus. Die Bedeutung der negativen Ionen besteht darin, daß sie während der Nacht die freien Elektronen speichern und bei Sonnenaufgang sehr schnell wieder freigeben.
Die freien *Elektronen* und Ionen im untersten Teil der Ionosphäre (< 62 km) sind nicht solaren, sondern kosmischen Ursprungs. Sie werden in allen Breiten durch die einfallende hochenergetische kosmische Strahlung erzeugt.

10.1.4. Physik der Magnetosphäre

Unter der Magnetosphäre versteht man das Gebiet des erdnahen Raums, in dem das Magnetfeld der Erde alle großräumigen Prozesse entscheidend mitbestimmt.
Oberhalb 100 km kommt die durch das Magnetfeld auf die geladenen Teilchen ausgeübte Kraft in die gleiche Größenordnung wie die Reibung der geladenen mit den Neutralteilchen. Hier liegt die untere Grenze der Magnetosphäre. Die äußere Grenze, die *Magnetopause*, auch magnetosheath genannt, (Abb. 10.1.4-1), grenzt den von dem Planeten Erde beeinflußten Raum gegen den interplanetaren Raum ab.
Die *Magnetosphärenphysik* beschäftigt sich mit den elektrischen und magnetischen Feldern und dem quasineutralen stoßfreien vollionisierten Plasma in diesem Gebiet sowie mit den Wechselwirkungen der Felder und Teilchen miteinander. In der Magnetosphäre laufen zahlreiche fundamentale plasmaphysikalische Prozesse ab, die aufgrund der erforderlichen extremen Bedingungen und großen Dimensionen im Labor nicht realisierbar sind, andererseits aber auch an anderen Orten im Universum, in der Sternatmosphäre, den Pulsaren und dem interstellaren Gas auftreten. Dazu gehören Stoßfronten, Instabilitäten, Teilchenbeschleunigungs-, Einfangs- und Diffusionsprozesse ebenso wie die Untersuchung von Strukturen an Grenzflächen zwischen stoßfreien Plasmen in verschiedenen Zuständen. Diese Nutzung der Magnetosphäre als „Großraumlabor" steht erst am Anfang.
10.1.4.1. Wechselwirkung Erdfeld – Sonnenwind. Definitionsgemäß muß zur Beschreibung der *Geometrie der Magnetosphäre* die Geometrie des Magnetfelds betrachtet werden. Diese wird durch die Wechselwirkung des erdmagnetischen Felds mit dem als Sonnenwind bezeichneten, ständig von der Sonne anströmenden Magnetoplasma bestimmt.

Der *Sonnenwind* ist ein quasineutrales, stoßfreies Plasma mit einer Dichte von 10^6 bis $5 \cdot 10^6$ Teilchen pro m^3, das einen Teil des solaren Magnetfelds der Korona mit sich führt. Im Mittel beträgt die magnetische Feldstärke 5 nT. Der Sonnenwind und das in ihm eingefrorene Magnetfeld strömen nicht längs der direkten Verbindung Sonne – Erde an, sondern im Mittel aus einer um 45° von dieser Richtung abweichenden Richtung. Das eingefrorene *interplanetare Magnetfeld* liegt nicht in der Ekliptikebene, sondern besitzt eine kleine Komponente senkrecht zur Ekliptikebene, die nord- oder südwärts gerichtet sein kann. Diese Komponente spielt für die Deutung der Energieübertragung durch Reconnectionsmodelle eine große Rolle. Die mittlere Geschwindigkeit des Sonnenwinds liegt bei 385 $km\,s^{-1}$, seine Bewegungsenergie beträgt einige 10^{-10} bis 10^{-9} Jm^{-3}, als thermische Energiedichte findet man $5 \cdot 10^{-12}$ bis $7 \cdot 10^{-12}$ Jm^{-3} und als Magnetfeldenergiedichte 10^{-11} Jm^{-3}.
Berechnet man die Alfvénsche Mach-Zahl, so findet man Werte zwischen 3 und 10, d. h. der Sonnenwind stellt eine magnetohydrodynamische „Überschallströmung" dar. Für den anströmenden Sonnenwind ist das Erdmagnetfeld ein Hindernis. Aufgrund der Überschallströmung kommt es zur Ausbildung einer *stehenden Stoßwelle im stoßfreien Plasma* im Abstand von einigen Erdradien vor der Magnetopause (↑ Abb. 10.1.4-1). Das Gebiet zwischen dieser Stoßwelle und der Magnetopause wird als *Übergangsregion*, häufig auch als *magnetosheath*, bezeichnet. Hier wird das Überschallplasma des Sonnenwinds auf Unterschallgeschwindigkeiten abgebremst. Magnetfeld- und Plasmamessungen zeigen einen sehr turbulenten, stark variierenden Charakter.
Bewegt man sich von der Sonne her weiter auf die Erde zu, so durchquert man bei etwa 10 Erdradien die *Magnetopause*. Abb. 10.1.4-2 zeigt, daß das Erdmagnetfeld zur Sonne hin durch den Sonnenwind stark zusammengedrückt wird; zur Nachtseite hin bildet sich dagegen ein lang gestreckter, bis zu mehr als 100 Erdradien langer *Magnetschweif*, den man sich aus Feldlinien gebildet vorstellen kann, die durch den Sonnenwind auf der Tagseite aufgerissen und auf die Nachtseite hin weggeblasen werden.
Um diese Beobachtungen physikalisch deuten zu können, muß das *Energieaustauschproblem Sonnenwind – Magnetosphäre* gelöst werden. Abschätzungen aufgrund von Beobachtungen zeigen, daß etwa 1 bis 2% der Energie des Sonnenwinds an die Magnetosphäre abgegeben werden muß. Von einem wirklichen Verständnis oder gar einer quantitativen Beschreibung dieser Energieaustauschmechanismen ist man weit entfernt. Qualitativ gibt es zwei Grundmodelle: das *Reconnectionsmodell* [*reconnection* engl.,

Abb. 10.1.4-2 Dreidimensionale Darstellung der Magnetosphäre nach HEIKKILLA und ROEDERER mit besonderer Beachtung der thermischen Plasmakomponente („polar cusp" polare Hörner)

Wiederverbindung], das ursprünglich von DUNGEY und PIDDINGTON entwickelt wurde, und das *Reibungsmodell* von AXFORD und HINES. Bei beiden Modellen handelt es sich um komplizierte, nichtlineare, plasmaphysikalische Prozesse, die quantitativ noch nicht beherrscht werden.

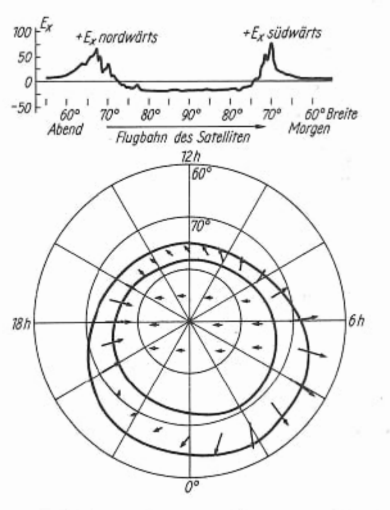


Abb. 10.1.4-4 Tageszeitliche und Breitenabhängigkeit des elektrischen Felds in der Ionosphäre in hohen Breiten

Grob anschaulich wird beim Axford-Hines-Modell an „unebenen, rauhen" Stellen der Magnetopause durch die Reibung des vorbeistreichenden Sonnenwinds Energie an das Plasma der Magnetosphäre abgegeben. Hier handelt es sich also um viele kleinräumige Prozesse.

Beim Reconnectionmodell verbinden sich die äußeren Feldlinien des Erdmagnetfelds mit den im Sonnenwind eingefrorenen interplanetaren Magnetfeldlinien; die Erdmagnetfeldlinien werden geöffnet oder auseinander geschnitten und mit dem Sonnenwind von der Sonne weg, d. h. die Flanken der Magnetopause entlang, zum Magnetschweif geführt. Im entfernten Magnetschweif tritt dann eine Wiedervereinigung der Magnetfeldlinienteile ein. Bei diesem Prozeß wird nicht nur Sonnenwindplasma mit in die Magnetosphäre hinein gebracht, sondern vor allem magnetische Energie in kinetische Energie umgewandelt. Besonders effektiv ist dieser Prozeß, wenn das interplanetare Magnetfeld eine südwärts zur Ekliptik geneigte Komponente besitzt. Die wiedervereinigten Magnetfeldlinien driften dann von der Nacht- zur Tagseite zurück, so daß hier ein ständig wirkender Prozeß aufgebaut werden könnte.

10.1.4.2. Magnetfelder, elektrische Felder und Ströme in der Magnetosphäre. Als Ergebnis der Wechselwirkung mit dem Sonnenwind entsteht auf der der Sonne zugewandten Seite eine Kompression des Magnetfelds, auf der Nachtseite dagegen der Magnetschweif, dessen Feldlinien auf der einen Seite von der Erde weg, auf der anderen Seite zur Erde hin gerichtet sind. Im Inneren des Schweifs muß sich daher eine Zone ausbilden, in der sich die Magnetfeldvektoren nahezu kompensieren bzw. neutralisieren, die *Neutralschicht* (↑ Abb. 10.1.4-2). Sie kann ungefähr 10 bis 12 Erdradien an die Erde heranreichen. Um ein Bild von der Stärke des Magnetfelds in der äußeren Magnetosphäre und im Schweif zu geben, zeigt Abb. 10.1.4-3a auf experimentellen Beobachtungen beruhende Linien konstanter magnetischer Feldstärke in der magnetischen Äquatorebene nach FAIRFIELD.

Die starke Abweichung vom Dipolfeld kann nur durch das Auftreten zusätzlicher Magnetfelder erklärt werden, die ihrerseits wiederum die Folge von Strömen sind. Um das beobachtete Magnetfeldbild zu erhalten, muß man daher auch in der Magnetopause und dem Gebiet der Neutralschicht komplizierte *Stromsysteme* annehmen. In Abb. 10.1.4-2 ist ein dreidimensionales Bild der Magnetosphäre nach HEIKKILA wiedergegeben, in dem als breite weiße Pfeile die erforderlichen Stromrichtungen eingezeichnet sind.

Zusätzliche Magnetfelder, der sogenannte Ringstrom, können auch im Inneren der Magnetosphäre durch größere Driftströme geladener Teilchen und durch die diamagnetische Wirkung geladener Teilchen entstehen. Wegen der außerordentlich hohen Parallelleitfähigkeit (↑ 10.1.3.2.) stellen die magnetischen Feldlinien im allgemeinen Äquipotentiallinien dar, d. h.,

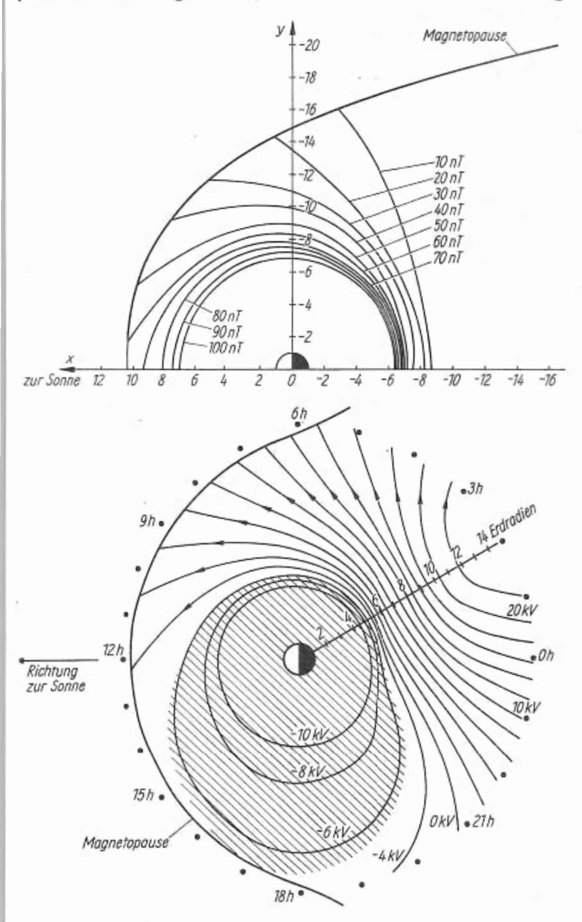


Abb. 10.1.4-3 a) Magnetfeldverlauf nach FAIRFIELD; b) Linien gleichen elektrischen Potentials in kV in der Äquatorebene in Abhängigkeit von der Ortszeit (LT) nach MCILLWAIN

in der Ionosphäre erzeugte elektrische Felder werden in die Magnetosphäre transportiert und ebenso in der Magnetosphäre erzeugte in die Ionosphäre.

Zahlreiche indirekte Beweise haben zu dem Schluß geführt, daß es außer den ionosphärisch bedingten dynamo-elektrischen Feldern (s. a. 10.1.3.2.) wenigstens eine, wahrscheinlich jedoch mehrere magnetosphärische Quellen für elektrische Felder gibt. Diese magnetosphärisch bedingten Felder werden durch ihre Auswirkungen auf die Bewegung geladener Teilchen erschlossen.

Abb. 10.1.4-3b zeigt den Verlauf eines so erschlossenen elektrischen Felds in der Äquatorebene in einem magnetosphärenfesten Koordinatensystem, in dem auch das durch die Rotation der Erde mit ihrem Magnetfeld hervorgerufene elektrische Korotationsfeld mit zu beachten ist.

Aus der Darstellung des Magnetfeldverlaufs (↑ Abb. 10.1.4-1, 10.1.4-2) erkennt man, daß die Äquipotentiallinien mit abnehmender Höhe immer dichter zusammenrücken, die elektrische Feldstärke also zunimmt und dadurch bessere Meßbedingungen entstehen. Magnetosphärisch bedingte elektrische Felder projizieren sich in die Ionosphäre in hohen Breiten. Abb. 10.1.4-4 zeigt Zusammenstellungen der in hohen Breiten erhaltenen Meßergebnisse. Man erkennt, daß es eine ovalförmige bis kreisrunde Zone starker elektrischer Feldstärke gibt, deren Zentrum nicht mit der Lage des geometrischen Pols zusammenfällt. Zu noch höheren Breiten hin, d. h. in der Polkappe, nimmt die Feldstärke dann wieder ab, ist aber stets so gerichtet, daß sich die geladenen Teilchen unter der Wirkung des elektrischen Felds von der Sonne weg bewegen. Da es sich um die Projektionen eines magnetosphärenfesten Koordinatensystems handelt, muß eine starke Abhängigkeit von der magnetischen Ortszeit erwartet werden. Diese ist für die elektrischen Felder dargestellt. Alle Messungen zeigen, daß in niederen und mittleren Breiten ionosphärisch bedingte elektrische Felder vorherrschen.

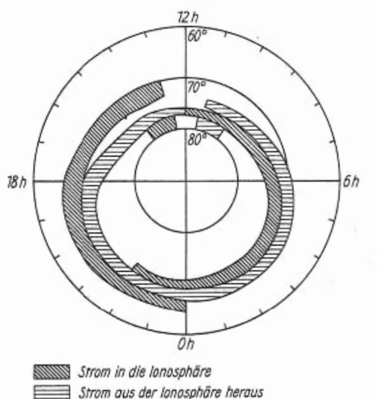


Abb. 10.1.4-5 Verteilung feldliniengerichteter Ströme in Abhängigkeit von Breite und Ortszeit nach Messungen des Satelliten TRIAD (Blick auf den Nordpol)

Im Gebiet hoher Breiten findet man parallel zu den magnetischen Kraftlinien verlaufende Stromschichten. Diese feldliniengerichteten Ströme existieren ständig (Abb. 10.1.4-5). Sie treten paarweise auf. Auf der Morgenseite fließt in der mehr zum Pol hin liegenden Stromschicht der Strom in die Ionosphäre hinein, in der mehr zu niederen Breiten hin liegenden Schicht aus der Ionosphäre heraus. Auf der Abendseite kehrt sich dieses Verhalten um. Die Stromstärke ist stark von der magnetischen Aktivität, d. h. von der Intensität der Wechselwirkung mit dem Sonnenwind abhängig. Auf der Tagesseite und um Mitternacht herum wird dieses Bild wesentlich komplizierter. Die feldliniengerichteten Ströme sind mit Sicherheit Teil eines großen umfassenden ionosphärisch-magnetosphärischen Stromsystems. Heute ist es jedoch weder möglich festzustellen, welche Teilchen Träger der Ströme sind, noch zu sagen, mit welchen magnetosphärischen Strömen die feldliniengerichteten Ströme verbunden sind.

10.1.4.3. Plasma in der Magnetosphäre. Die gesamte Magnetosphäre ist mit einem thermischen, niederenergetischen ($E < 1$ keV), quasineutralen und stark verdünnten, im wesentlichen aus Elektronen und Protonen bestehenden Plasma erfüllt.

Neben der Stoßfreiheit ist dies das hervorstechendste Merkmal der Magnetosphäre. Das *thermische Hintergrundplasma* befindet sich häufig im Zustand lokaler Gleichgewichte.

In einigen Gebieten der Magnetosphäre findet man außerdem eine Plasmakomponente höherer Energie (1 keV $\leq E \leq$ 100 keV). Die Dichte der höherenergetischen Komponente ist um zwei bis drei Größenordnungen niedriger als die des Hintergrundplasmas. Hochenergetische Partikeln mit $E >$ 100 keV sind vorwiegend in den sogenannten Strahlungsgürteln konzentriert.

Wenn man annehmen kann, daß das Hintergrundplasma mit den höherenergetischen Plasmakomponenten nicht in unmittelbarer Wechselwirkung steht, so kann man beide Plasmabestandteile als unabhängig voneinander betrachten. Das erleichtert die Beschreibung des Plasmaverhaltens in den vorgegebenen Feldkonfigurationen der magnetischen und lektrischen Felder wesentlich.

Betrachtet man das Hintergrund- und niederenergetische Plasma, so kann man grob zwischen Plasmasphäre, Trogregion, Plasmaschicht im Schweif, Plasmamantel oder Eintrittszone, polaren Hörnern und Randschicht in niederen

Breiten unterscheiden (↑ Abb. 10.1.4-1,10.1.4-2). Tab. 10.1.4-6 faßt typische Größen des Hintergrundplasmas einiger dieser Gebiete zusammen. Das niederenergetische Plasma unterliegt vorwiegend einer von der Ladung und Energie der Teilchen unabhängigen Drift in gekreuzten elektrischen und magnetischen Feldern.

Das *Plasma in der Plasmasphäre* rotiert unter der Wirkung des Korotationsfelds mit der Erde, die Lage der äußeren Grenze der Plasmasphäre, der Plasmapause, ist identisch mit der Lage der letzten geschlossenen Äquipotentiallinie (↑ Abb. 10.1.4-3 b). Man nimmt an, daß das niederenergetische Plasma der Plasmasphäre aus der Ionosphäre stammt und weiter aufgeheizt wurde. Unmittelbar an die Plasmapause schließt sich der durch eine stark verminderte Teilchendichte gekennzeichnete Plasma- oder Elektronentrog an (↑ Abb. 10.1.4-2).

Die *Plasmaschicht im Schweif* (↑ Abb. 10.1.4-1, 10.1.4-2) stellt eine symmetrisch um die Neutralschicht liegende 6 bis 10 Erdradien dicke Schicht hoher Flüsse niederenergetischer Elektronen und Protonen dar. Die thermische Energie kann hier 1 keV betragen, die Dichte 10^5 bis 10^6 m^{-3}. Der breite erdnahe etwa 0,5 Erdradien dicke innere Rand der Plasmaschicht wird durch einen plötzlichen exponentiellen Anstieg der Elektronenenergiedichte charakterisiert. Seine Lage hängt von der Teilchenart und Teilchenenergie ab, wobei Protonen etwa um einen Erdradius tiefer in die Magnetosphäre eindringen können. Man nimmt an, daß es sich beim Plasma in der Plasmaschicht vorrangig um Sonnenwindpartikel handelt, die über den entfernten Schweif oder durch die polaren Hörner in die Magnetosphäre eingedrungen sind.

Eine wichtige Entdeckung der letzten Jahre besteht im Auffinden der *polaren Hörnerbänder* [engl. *polar cusps*] (↑ Abb. 10.1.4-1, 10.1.4-2). Durch sie kann niederenergetisches Plasma aus der Übergangsregion in die Magnetosphäre eintreten und dann zum Schweif hin driften und die Plasmaschicht füllen. Auf der Magnetopause haben die Cusp-Gebiete eine Breite von ungefähr einem Erdradius, ihr ionosphärischer Fußpunkt liegt bei 70° bis 80° invarianter Breite. In den letzten Jahren fand man in unmittelbarer Nähe der Magnetopause Bereiche, in denen das niederenergetische Plasma eindeutig durch die Wechselwirkung mit dem Sonnenwind mitbestimmt wird. Hierzu gehören der Plasmamantel, die Eintrittsschicht und die Randschicht in niederen Breiten.

Eine unabhängige getrennte Behandlung des Hintergrund- und des hochenergetischen Plasmas ist nicht mehr möglich, wenn eine enge Kopplung zwischen beiden vorliegt, wenn also z. B. das Hintergrundplasma die räumlich-zeitlich variablen elektromagnetischen Felder liefert, in denen sich die energiereichen Teilchen bewegen oder wenn es zu Resonanzen zwischen Wellen im Hintergrundplasma und Oszillationsbewegungen der höherenergetischen Teilchen kommt.

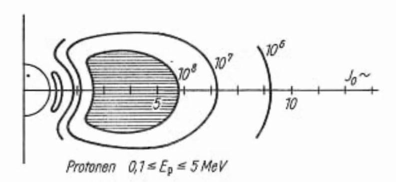
Besteht zwischen beiden Plasmakomponenten nur eine schwache Wechselwirkung, so können in vielen Fällen die Eigenschaften hochenergetischer Teilchen mit Hilfe einer Einteilchen- oder Testpartikeltheorie, der Orbittheorie [*orbit* engl., Bahn], beschrieben werden.

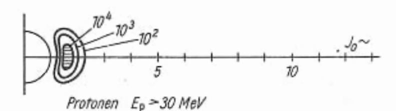
Die *hochenergetische Komponente des Plasmas* zeigt eine von der des Hintergrundplasmas vollkommen abweichende Verteilung, was sowohl auf die andersartigen Einspeisungs- und Erzeugungsmechanismen als auch auf die Tatsache

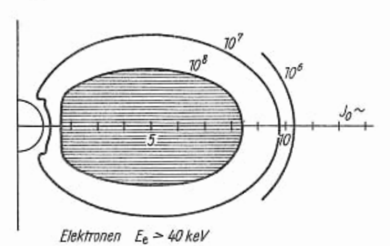
Tab. 10.1.4-6 Plasmaparameter des thermischen Hintergrundplasmas

	Plasmasphäre	Plasmatrog	Plasmaschicht
Dichte n in m^{-3}	10^9	10^6	10^5–10^6
Temperatur T in K	$5 \cdot 10^3$	$5 \cdot 10^3$	10^7
Magnetfeld B in Wb m^{-2}	10^{-6}	10^{-7}	$(1–5) \cdot 10^{-7}$
$\beta = \frac{Tkn}{B^2/2\mu_0}$	10^{-4}	10^{-2}	10^0–10^{-2}
thermische Geschwindigkeit der Elektronen in m s^{-1}	$3 \cdot 10^5$	$3 \cdot 10^5$	10^7
thermische Geschwindigkeit der Ionen in m s^{-1}	10^4	10^4	?
Plasmafrequenz der Elektronen ω_{pe} in s^{-1}	10^6	$3 \cdot 10^4$	10^3
Plasmafrequenz der Ionen ω in s^{-1}	$2 \cdot 10^4$	$6 \cdot 10^2$	1
Debyeradius in m	0,3	~ 3	$\sim 3 \cdot 10^2$
Gyrofrequenz der Elektronen ω in s^{-1}	$\sim 10^5$	$\sim 10^4$	10^3
Gyrofrequenz der Ionen ω in s^{-1}	50	10	10^{-1}

zurückzuführen ist, daß die Wirkung elektrischer Felder gegenüber der des Magnetfelds mit zunehmender Energie der Teilchen schnell abnimmt. Hochenergetische Teilchen folgen bei ihrer Drift Linien konstanter Magnetfeldstärke, Protonen driften westwärts, Elektronen ostwärts. Die in der Abb. 10.1.4-3a gezeigten Linien konstanter Magnetfeldstärke können







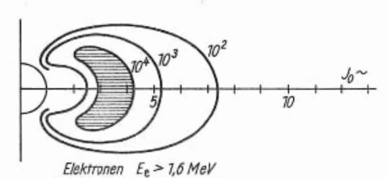


Abb. 10.1.4-7 Linien konstanter Flüsse J_0 in $cm^{-2}\,s^{-1}$ für hochenergetische Protonen und Elektronen in den Strahlungsgürteln nach VAN ALLEN

daher auch als Driftbahnen hochenergetischer Partikel betrachtet werden. Es sind zwei verschiedene Gebiete zu erkennen: Geschlossene Linien konstanter Magnetfeldstärke, die die Magnetopause nicht berühren, stellen das Gebiet stabil eingefangener Teilchen dar. In diesem Gebiet können hochenergetische Teilchen eine volle Drift um die Erde herum ausführen. Für Teilchen auf Bahnen, die die Magnetopause berühren, ist ein Umdriften der Erde nicht möglich, sie befinden sich im Gebiet zeitweise eingefangener Teilchen.

Das Gebiet stabil eingefangener Teilchen wird als *Strahlungsgürtel* bezeichnet. Abb. 10.1.4-7 gibt einen Einblick in die räumliche Struktur und die Intensität von Strahlungsgürtel-Elektronen und -Protonen. Man erkennt, daß der Schwerpunkt der räumlichen Verteilung höherenergetischer Teilchen dichter zur Erde hin liegt. Es wird zwischen innerem und äußerem Strahlungsgürtel unterschieden. Der innere Strahlungsgürtel wird vor allem durch sehr hochenergetische Protonen, der äußere vorwiegend durch hochenergetische Elektronen gebildet. Während die Intensität des inneren Strahlungsgürtels sehr geringe zeitliche Schwankungen zeigt, sind im äußeren Strahlungsgürtel wesentlich schnellere und stärkere Variationen zu beobachten, die vor allem im Zusammenhang mit geomagnetischen Stürmen auftreten. Die äußere Grenze des Gebiets stabil eingefangener Teilchen liegt auf der Nachtseite bei etwa fünf bis sechs Erdradien.

Der Fluß hochenergetischer Partikel nimmt außerhalb der Strahlungsgürtel um etwa zwei Größenordnungen ab und unterliegt starken räumlichen und zeitlichen Änderungen. Die hochenergetischen Partikel werden vermutlich durch in der Magnetosphäre ablaufende Beschleunigungsprozesse aus Sonnenwind- oder Atmosphärenteilchen erzeugt.

Besonders interessante Verhältnisse liegen vor, wenn die Annahmen der einfachen Orbittheorie teilweise oder völlig verletzt sind, d. h., wenn es zu starken Wechselwirkungen kommt. Für eine ganze Reihe derartiger Prozesse liegen Theorien vor. Diese Vorgänge spielen sowohl für die Beschleunigung von Partikeln in der Magnetosphäre als auch für die Deutung der Prezipitation von Teilchen eine Rolle. Unter *Prezipitation* versteht man das Ausscheiden des Teilchens aus der eingefangenen Verteilung durch Vordringen in dichtere Schichten und die Abgabe seiner Energie an die dort vorhandenen neutralen oder geladenen Teilchen. Die Prezipitation von Teilchen stellt ebenso wie die feldliniengerichteten Ströme eine der Seiten der Kopplung der Magnetosphäre mit der Ionosphäre dar.

10.2. Erscheinungsformen der Materie im Sonnensystem

10.2.1 Mond

10.2.1.1. Mondbahn und Mondbewegung. Der Mond ist der der Erde am nächsten stehende Himmelskörper. Seine mittlere Entfernung vom Erdmittelpunkt beträgt 384401 km. Der erdnächste Punkt, das *Perigäum*, ist 356334 km, der erdfernste, das *Apogäum*, 406610 km entfernt. Die Bahnebenen von Erde und Mond bilden einen Winkel von 5°9′.

Die Zeitspanne zwischen zwei aufeinander-

folgenden Durchgängen des Monds durch einen bestimmten Bahnpunkt definiert den *Monat*. Da die Erde von der Kugelgestalt abweicht und neben ihrer Anziehungskraft auf den Mond auch die der Sonne nicht vernachlässigt werden kann, führt der Erdtrabant eine sehr komplizierte Bahnbewegung aus. Die Schnittgröße zwischen den Bahnebenen von Mond und Erde, die *Knotenlinie*, dreht sich entgegengesetzt zur Bewegungsrichtung des Monds, während die Verbindungsgerade von Perigäum und Apogäum, die *Apsidenlinie*, in Bewegungsrichtung wandert. Benutzt man daher die Knoten oder die Apsiden der Mondbahn als Bezugspunkte, so ergeben sich verschieden lange Monate (Tab. 10.2.1-1).
Neben periodischen Änderungen wie den eben genannten, gibt es auch nichtperiodische Störungen der Mondbewegung, z. B. die *säkulare Akzeleration*, durch die der Mond pro Jahrhundert um 0″,08 seiner theoretisch berechneten Position vorausläuft und sich damit allmählich von der Erde wegbewegt.

10.2.1.2. Physikalische Eigenschaften des Monds. Der *Durchmesser* des Monds beträgt 3476 km. Während die sichtbare Mondscheibe fast ideale Kreisform zeigt, ist der Monddurchmesser längs der Verbindungslinie Erde – Mond um etwa 4 km größer als der obige Wert. Diese Zahl übersteigt die durch die Gezeitenwirkung der Erde auf den starren Mond zu erwartende Verformung. Der Mond dreht sich in der gleichen Zeit einmal um seine Achse, in der er die Erde umrundet. Durch diese *gebundene Rotation* wendet er der Erde immer die gleiche Seite zu.
Die *Masse* des Monds beträgt $7{,}348 \cdot 10^{22}$ kg, d. h. 1,23% der Erdmasse. Damit weist der Erdtrabant ein ungewöhnlich hohes Massenverhältnis zu einem Planeten auf. Die mittlere *Dichte* des Monds ist mit 3341 kgm^{-3} wesentlich kleiner als die der Erde. Die *Schwerebeschleunigung* an der Mondoberfläche beträgt nur 16,5% der der Erde. Die *Fluchtgeschwindigkeit* von 2,38 kms^{-1} ist so niedrig, daß aus dem Mondinneren freigesetztes Gas in verhältnismäßig kurzer Zeit in den Weltraum entweichen kann. Der Mond ist daher ein atmosphäreloser Himmelskörper. Auf seiner Tagseite heizt sich der dunkle Mondboden auf *Temperaturen* bis zu 390 K auf, während in der Mondnacht die Temperatur auf 120 K fallen kann.
Der Mond besitzt eine starre Gesteinshülle. Die oberste Schicht dieser bis in 1000 km Tiefe reichenden festen Lithosphäre, die *Mondkruste*, ist auf der erdzugewandten Seite 50 bis 65 km dick, auf der Rückseite ist ihre Dicke annähernd doppelt so groß. Unter der Kruste befindet sich der aus schwererem Gestein bestehende *Mondmantel*. Wegen der Starrheit seiner Lithosphäre. herrscht auf dem Mond, wie dort installierte Seismometer zeigten, weitgehend seismische Stille. Die wenigen echten *Mondbeben* gehen auf Herde in großer Tiefe zurück und treten vor

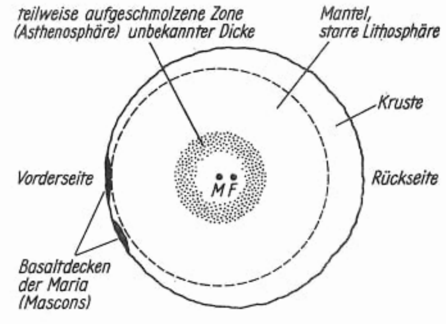


Abb. 10.2.1-2 Innerer Aufbau des Monds; Krustendicke ist übertrieben eingezeichnet; *M* Massenzentrum, *F* geometrisches Zentrum des Monds

Tab. 10.2.1-1 Länge der Monate in mittlerer Sonnenzeit

Monat	Bezugspunkt in der Mondbahn	Monatslänge d	h	min	s	Ursache der Abweichung vom siderischen Monat
siderischer	Stundenkreis eines Fixsterns	27	7	43	11,5	
synodischer	Stundenkreis der Sonne	29	12	44	2,9	Bewegung der Erde um die Sonne
tropischer	Stundenkreis des Frühlingspunktes	27	7	43	4,7	Präzession der Erdachse
drakonitischer	aufsteigender Knoten	27	5	5	35,8	Drehung der Knotenlinie
anomalistischer	Perigäum	27	13	18	33,2	Drehung der Apsidenlinie

allem dann auf, wenn der Mond sein Perigäum passiert.
Ob der *Mondkern* gegenwärtig noch flüssig ist und wie der Erdkern aus Eisen besteht, ist noch unklar. Die Magnetisierung des Mondgesteins deutet an, daß zumindest früher ein flüssiger Kern aus elektrisch leitendem Material vorhanden war und für ein globales Magnetfeld sorgte (Abb. 10.2.1-2).

10.2.1.3. Geologische Beschaffenheit der Mondoberfläche. Bereits mit bloßem Auge sind auf dem Mond helle und dunkle Flächen erkennbar. Erstere nehmen rund 2/3 der sichtbaren Mondhemisphäre ein. Diese höher liegenden, zerklüfteten und dicht mit *Kratern* bedeckten Flächen werden *Hochländer*, *Kontinente* oder *Terraflächen* genannt. Die dunklen Gebiete sind an großen Kratern arme Tiefebenen, die man aus historischen Gründen *Mondmeere*, *Mareflächen* oder einfach *Maria* [lat. Meere, Pl. zu Mare] nennt. Von kleinen Gebieten abgesehen, hat die *Mondrückseite* kontinentalen Charakter und ist dicht mit Kratern bedeckt. Wie auf der Vorderseite tragen auch die größeren Krater der Rückseite inzwischen die Namen bedeutender Gelehrter.
Während das Alter der Terraflächen zwischen 4 und 4,6 Mrd. Jahren liegt, bildeten sich die Maria vor 3,9 bis 3,1 Mrd. Jahren, als basaltische Lava aus dem Mondmantel emporquoll und die Niederungen der Mondoberfläche füllte. Maria wie z. B. Mare Imbrium, Mare Serenitatis oder Mare Crisium, die nahezu kreisförmig begrenzt sind und daher *Zirkularmaria* genannt werden, sind Becken von Riesenkratern der ursprünglichen Mondkruste, die sich in dieser Zeit mit Lava füllten. Durch die Anhäufung schweren Gesteins in diesen Becken kam es zur Ausbildung von Massenkonzentrationen, *Mascons*, die als lokale Schwereanomalien auffallen. Der *Mondvulkanismus* z. Z. der Marebildung führte primär zu mächtigen flächenhaften Ergüssen und erzeugte nur sekundär bescheidene Vulkanbauten, z. B. die *Kuppeln* oder *Dome*, die erst auf den Bildern von Mondsatelliten entdeckt wurden. Typisch für die Mondmeere sind wie Runzeln aussehende Höhenrücken und an Flußläufe erinnernde gewundene Rillen, die sogenannten *Sinusrillen*, die wahrscheinlich Lavastromtäler oder eingebrochene Lavatunnel darstellen. Mit Ausnahme von gelegentlichen *Gasausbrüchen* gibt es heute auf dem Mond keinen Vulkanismus mehr.
Vor der Marebildung war die Mondkruste einem intensiven, aber rasch abklingenden *Meteoritenbombardement* ausgesetzt. Dabei entstand die für die Terraflächen typische Sättigung mit Einschlags- oder *Impaktkratern*. Die in den Mareflächen gelegenen großen Krater, z. B. Kopernikus, oder junge Impaktkrater auf den Mondterra, z. B. Tycho, stammen von vereinzelten Nachzüglern und fallen durch das frische Aussehen ihres Walls und der beim Einschlag ausgeworfenen Trümmerdecke und durch ein radiales System heller Strahlen auf (Abb. 10.2.1-3).
Da der Mond nie eine Atmosphäre besaß, gab und gibt es keine Verwitterung, keine Erosion durch Wasser und Wind, keinen Materialtransport und keine Sedimentation.

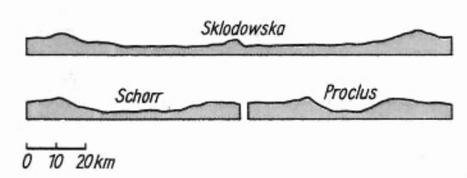


Abb. 10.2.1-3 Charakteristische Profile von Mondkratern

Das *Altern* der Mondoberfläche, z. B. der Kraterwälle, wurde und wird nur durch Meteoriteneinschläge und durch deren *tektonische* und *vulkanische Folgeerscheinungen* bewirkt. Auf dem Mond gibt es keine Faltengebirge. Bekannte *Mondgebirge*, wie Alpen, Apenninen und Kaukasus, sind Krustenschollen, die bei der Bildung des Imbriumbeckens, aus dem später das Mare Imbrium hervorging, als Kraterwall aufgetürmt wurden.
Wegen der schnellen und tiefgreifenden Erstarrung findet man auf dem Mond im Unterschied zur Erde keine Anzeichen für driftende lithosphärische Platten. Die auf dem Mond vorhandenen *tektonischen Bildungen*, z. B. das Alpenquertal, die Gerade Wand im Mare Nubium oder die zahlreichen geraden Rillen und Spalten, sind wahrscheinlich ziemlich alte Strukturen. Die tektonische Aktivität der Mondkruste erlosch genauso schnell wie die vulkanische. Geologisch ist der Mond als toter Himmelskörper einzustufen (Tafel 53).

10.2.1.4. Eigenschaften des Mondgesteins. Als Folge von Impakt und Kraterbildung ist die gesamte Mondoberfläche von einer 6 bis 12 m dicken Trümmerschicht bedeckt. Dieser *Regolith* besteht aus staubfeinem Material, aus Gesteins- und Mineralfragmenten und zahlreichen Glaskügelchen, die als Kondensationsprodukte verdampften Gesteins aufzufassen sind. In dieses feine Material sind zahlreiche Steine aller Größen eingebettet. Sie bestehen häufig nicht aus einheitlichem kristallinem Gestein, sondern aus *Brekzien*, d. h. mechanisch verfestigten Gemischen aus Gesteinskörnchen verschiedener Herkunft. Man kann sie als durch Impakt zusammengeschweißten Regolith auffassen (Tafel 53).
Das Mondgestein ist generell älter, wesentlich mineralärmer und viel einheitlicher zusammengesetzt als das irdische. Es kommen nur *basische Gesteine* vor, also solche mit Silicaten, die arm

an SiO_2 sind. Wasserhaltige Minerale fehlen gänzlich. Das typische Maregestein sind dunkle *Basalte*. Auf den Terraflächen dominieren Brekzien aus Gesteinen wie *Anorthosit*, *Norit* und *Troktolit*. Es gibt dort aber auch aluminiumreiche Basalte. Eine Gruppe von Basalten, Noriten und hellen Brekzien mit einem auffällig hohen Gehalt an radioaktiven Elementen faßt man unter der Bezeichnung *Kreep* zusammen, einem Kunstwort, das wegen des Gehalts an *K*alium, Seltenen Erden [engl. *r*are *e*arth *e*lements] und *P*hosphor geprägt wurde (Tab. 10.2.1-4).

Tab. 10.2.1-4 Chemische Zusammensetzung des Mondgesteins nach Wood in Masse-%

	Marebasalt Apollo 11	Luna 16	Anorthosit Apollo 16	Troktolith Apollo 16	Kreepnorit Apollo 14
SiO_2	41,00	45,50	43,97	45,16	47,90
Al_2O_3	9,83	13,95	35,83	20,05	15,60
TiO_2	10,28	4,04	0,02	0,70	1,80
MgO	7,77	5,95	0,25	14,85	10,93
FeO	19,03	17,77	0,36	6,40	10,74
CaO	11,96	11,96	18,95	11,85	9,90
Na_2O	0,37	0,63	0,34	0,48	0,78
K_2O	0,05	0,21	0,01	0,11	0,52
P_2O_5	0,07	0,15	0,00	0,15	0,60

Die Bestandteile aller Mondgesteine sind ursprünglich aus einer Schmelze auskristallisiert. Der Mond besteht also nicht aus primitivem meteoritischem Material, sondern seine Gesteine sind wie das primäre Gestein unseres Planeten in einem *magmatischen Differenzierungsprozeß* entstanden.

10.2.1.5. Entstehung des Monds. Trotz großer Fortschritte der Mondforschung durch die Raumfahrt ist über die Entstehung des Monds nichts Sicheres bekannt. Die drei klassischen Hypothesen zur Mondentstehung sind:

Der Mond ging durch *Spaltung* der Urerde in zwei Teile aus dieser hervor.

Erde und Mond entstanden als *Doppelplanet* gleichzeitig.

Der Mond entstand unabhängig von der Erde und wurde später von dieser *eingefangen*.

Davon läßt sich die erste wahrscheinlich ausschließen, denn die chemische Zusammensetzung des Mondgesteins, insbesondere das völlige Fehlen von Wasser im Mineralbestand der Mondkruste, weist auf gänzlich andere chemische Ausgangsbedingungen des Monds gegenüber der Erde hin. Für die dritte Hypothese ergibt sich nach der Kenntnis des Mondgesteins die Einschränkung, daß der Mond nicht in Bereichen außerhalb der Erdbahn entstanden sein kann, beispielsweise nicht im Planetoidengürtel.

10.2.2. Sonne

10.2.2.1. Sonne als Ganzes. Die Sonne ist der Zentralkörper des Planetensystems und im Mittel $1{,}4959787 \cdot 10^{11}$ m von der Erde entfernt. Sie ist ein gewöhnlicher Stern der Spektralklasse G 2 V (↑ 10.3.1.8.), dessen Energiestrom das Leben auf der Erde ermöglicht.

Die Sonne ist der einzige Stern, der wegen seiner großen Nähe detailliert erforscht werden kann.

Der *Sonnendurchmesser* beträgt $1{,}392 \cdot 10^9$ m, und ihre *Masse* ist mit $1{,}989 \cdot 10^{30}$ kg rund 333000mal so groß wie die der Erde bzw. 750mal so groß wie die aller anderen Körper des Sonnensystems zusammen. Die mittlere *Dichte* der Sonne beträgt 1408 kgm^{-3}, die *Schwerebeschleunigung* an der Oberfläche ist 27,9mal so groß wie die irdische. Die Sonne strahlt ständig $3{,}90 \cdot 10^{26}$ W aus. Der Strahlungsstrom am Ort der Erde, den man als *Solarkonstante* bezeichnet, beträgt 1387 Wm^{-2}. Entsprechend ihrer Strahlungsleistung beträgt die *effektive Temperatur* der Sonnenoberfläche 5800 K.

Die Sonne rotiert nicht wie ein starrer Körper. Die *siderische Rotationsdauer* beträgt am Äquator 25,03 d, das entspricht einer Rotationsgeschwindigkeit von2,02 kms^{-1}. Mit zunehmender heliographischer Breite wächst die Rotationsdauer. Sie beträgt bei 80° Breite rund 31 d. Die Sonne besitzt ein *Magnetfeld*, dessen Stärke wahrscheinlich nur 10^{-4} T beträgt. Ihm überlagern sich jedoch die um Größenordnungen stärkeren lokalen Magnetfelder in den Aktivitätsgebieten (↑ 10.2.2.3.).

10.2.2.2. Aufbau der Sonne. Die Strahlung, die wir von der Sonne empfangen, stammt aus ihren äußeren Schichten. Obgleich sich das *Sonneninnere* der unmittelbaren Beobachtung entzieht, kann sein Aufbau aus bestimmten Gleichgewichtsbedingungen unter Beachtung der Größen, die die Sonne als Ganzes charakterisieren, modellmäßig beschrieben werden. Man muß

dabei allerdings die Prozesse, denen die Sonne ihre gewaltige Energieproduktion verdankt, die Wechselwirkungsprozesse zwischen Stoff und Strahlung im Sonneninneren und die *Zustandsgleichung* des Sonnengases, d. h. den Zusammenhang zwischen Druck, Dichte und Temperatur, sehr genau kennen. Die Zustandsgleichung bereitet dabei die geringsten Probleme, weil das Gas im Sonneninneren vollständig ionisiert ist und sich wie ein ideales Plasma verhält. Aus der Berechnung solcher *Sonnenmodelle* geht hervor, daß Dichte und Temperatur zum Sonnenzentrum hin stark ansteigen. Im Zentrum beträgt die Temperatur $16 \cdot 10^6$ K, während die Dichte bei 150000 kgm^{-3} liegt. Wie die Spektralanalyse des von der Sonne ausgestrahlten Lichts zeigt, besteht das Sonnengas zu 73 Massen-% aus Wasserstoff, zu 26 Massen-% aus Helium und zu reichlich einem Massen-% aus schwereren Elementen. Unter den extremen Bedingungen des Sonneninneren finden *Kernfusionen* statt, bei denen aus Wasserstoff Helium aufgebaut wird. Diese Fusionen sind die Quelle der Sonnenenergie. Im Sonnenzentrum ist die Heliumhäufigkeit mehr als doppelt so groß wie in den Gebieten, in denen keine Kernfusionen stattfinden. Die freigesetzte Kernenergie wird im allgemeinen als Strahlung durch die Sonne nach außen transportiert. Infolge der abnehmenden Temperatur verschiebt sich dabei die spektrale Zusammensetzung der Strahlung zu immer größeren Wellenlängen. Handelt es sich im Sonnenkern um kurzwellige Röntgenstrahlung, so verläßt diese Energie die Sonne hauptsächlich als sichtbares Licht. Dicht unter der Sonnenoberfläche erfolgt der Energietransport durch Konvektion, weil hier der Wasser-

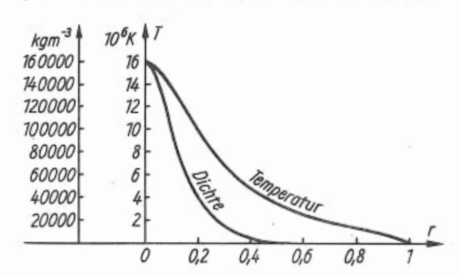


Abb. 10.2.2-1 Verlauf von Dichte ϱ und Temperatur T im Sonneninneren; Zentrumsabstand r in Einheiten des Sonnenradius

stoff rekombiniert und im neutralen Zustand stark Strahlung absorbiert. Diese *Wasserstoffkonvektionszone* hat eine Dicke von 10 bis 20% des Sonnenradius (Abb. 10.2.2-1).

Jene äußeren Bereiche der Sonne, aus denen die Strahlung direkt nach außen tritt, nennt man die *Sonnenatmosphäre*. Ihre unterste, knapp 500 km dicke Schicht heißt *Photosphäre*, denn aus ihr stammt fast das gesamte Sonnenlicht.

Die aus der Photosphäre austretende Strahlung hat ein kontinuierliches Spektrum, in dem die nach FRAUNHOFER benannten Absorptionslinien auftreten. Bisher fand man im Bereich des sichtbaren Lichts mehr als 20000 *Fraunhofersche Linien*, die zu etwa 73% identifiziert und 63 Elementen zugeordnet werden konnten. Die meisten stammen von Metallen, besonders vom Eisen.

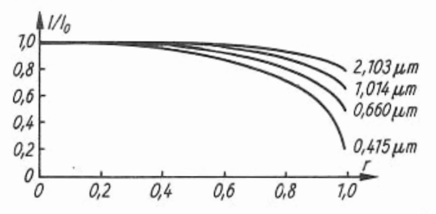


Abb. 10.2.2-2 Intensitätsverteilung über die Sonnenscheibe bei verschiedenen Wellenlängen; I_0 Intensität der Scheibenmitte, r Abstand von der Scheibenmitte in Einheiten des Sonnenradius

Sowohl im integralen Licht als auch bei den Fraunhoferschen Linien des Sonnenspektrums zeigen sich Unterschiede, je nachdem, ob man Licht von der Mitte oder vom Rand der Sonnenscheibe untersucht. Aus der im integralen Licht zu beobachtenden *Randverdunklung* kann man auf die Temperaturschichtung in der Photosphäre schließen. Die Temperatur beträgt in der oberen Photosphäre etwa 4500 K und steigt bis auf 7000 K am Boden der Photosphäre an. Dort beträgt die Gasdichte etwa 10^{-10} kgm^{-3} (Abb. 10.2.2-2).

Die Photosphäre erscheint nicht gleichmäßig hell, sondern zeigt eine als *Granulation* bezeichnete körnige Struktur. Auf dunklerem Hintergrund sind wabenförmige begrenzte, kleine, helle Gebilde, die *Granulen*, zu sehen, deren Durchmesser rund 1000 km beträgt (Tafel 54). Offensichtlich handelt es sich dabei um Konvektionszellen, deren Lebensdauer bei 5 bis 8 min liegt. Neben der Granulation gibt es noch ein gröberes Netz von Konvektionszellen mit Durchmessern zwischen 15000 und 40000 km, die *Supergranulation*, deren Lebensdauer im Durchschnitt bei 20 h liegt.

An die Photosphäre schließt sich die rund 2000 km dicke *Chromosphäre* an, in der die Gasdichte auf Werte um 10^{-15} kgm^{-3} gefallen ist. Ohne Hilfsmittel ist sie nur bei totalen Sonnenfinsternissen direkt zu beobachten. Wenn der Mond die Photosphäre gerade abdeckt, wird sie als farbiger Saum sichtbar. Ihr Spektrum zeigt die im Fraunhoferspektrum der Photosphäre vorhandenen Linien in Emssion. Von

der Chromosphäre schießen mit Geschwindigkeiten von 10 bis 30 kms^{-1} flammenähnliche Lichtzungen, die *Spiculen*, bis zu einer Höhe von 10000 km empor. Sie haben im Mittel Durchmesser von 1000 km und Lebensdauern von wenigen Minuten. Die Spiculen sind zwar chromosphärische Erscheinungen, ragen aber bereits in die nächste Schicht der Sonnenatmosphäre, in die Korona, hinein.

Zur Untersuchung des physikalischen Zustands der Chromosphäre benutzt man vor allem *Spektroheliogramme*. Das sind Sonnenaufnahmen im Licht eines extrem schmalen Wellenlängenbereichs, den man in markante Fraunhofersche Linien legt. Das dort beobachtete Licht stammt aus der Chromosphäre, und zwar aus einer um so höheren Schicht, je näher der ausgesonderte Spektralbereich am Zentrum der Fraunhoferschen Linie liegt. Auf den Spektroheliogrammen der hohen chromosphärischen Schichten beobachtet man ein helldunkles Maschenwerk, dessen helle Elemente *Flocculi* genannt werden.

Wegen der komplizierten physikalischen Bedingungen in der Chromosphäre gibt es noch keine vollständige Theorie über sie. Das Gas befindet sich hier nicht im thermischen Gleichgewicht. Die Temperatur steigt von 4500 K an ihrem unteren Rand auf einige 100000 K in der Höhe der Spiculen. Die Aufheizung der oberen Chromosphäre erfolgt durch Stoßwellen, deren Energie aus der unter der Photosphäre liegenden Wasserstoffkonvektionszone stammt und die Photosphäre als Schallwellen durchdringt. Diese erreichen wegen der geringeren Dichte in der Chromosphäre Überschallgeschwindigkeit und gehen zu Stoßwellen über. Die beobachteten Granulen und Spiculen hängen vermutlich mit diesen Vorgängen zusammen.

Die äußerste Schicht der Sonnenatmosphäre, die *Korona*, läßt sich bis zum Mehrfachen des Sonnendurchmessers verfolgen. In Form des Sonnenwinds erstrecken sich die Ausläufer der Korona über den gesamten Bereich des Planetensystems. Die Temperatur der Korona ist von der Größenordnung 10^6 K, so daß schwere Elemente wie Calcium, Nickel und Eisen in hohen Ionisationsstufen auftreten. Wegen dieses Umstands bildeten die Emissionslinien der Korona jahrzehntelang eines der hartnäckigsten Rätsel der Spektroskopie. Die Energie der Korona stammt aus den erwähnten Stoßwellen. Bereits 1/10000 der in den Granulen enthaltenen Energie reicht aus, um das Koronagas ständig a seiner hohen Temperatur zu halten. Wege.. dieser hohen Temperatur strömt ständig heißes Plasma mit eingefrorenen Magnetfeldern aus der Korona in den interplanetaren Raum und bildet den *Sonnenwind* (↑ 10.2.3.14.).

Das *Koronaspektrum* läßt sich in drei Komponenten zerlegen. Die *K-Komponente* zeigt ein rein kontinuierliches Spektrum, das dem der Photosphäre gleicht, aber keine Fraunhoferschen Linien enthält. Sie entsteht durch Streuung des photosphärischen Lichts an den Elektronen des Koronaplasmas, die sich wegen der hohen Temperatur so schnell bewegen, daß die Absorptionslinien durch den Dopplereffekt verwischt werden. In der *F-Komponente* treten die Fraunhoferschen Linien auf. Hier handelt es sich um das Streulicht der sonnennahen, zur Zodiakalwolke gehörenden Mikrometeorite

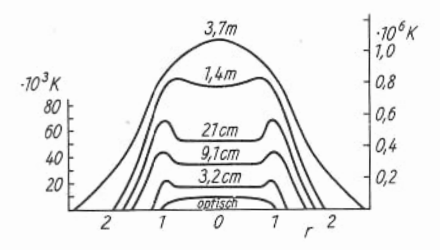


Abb. 10.2.2-3 Intensität (ausgedrückt in Äquivalenttemperatur) der Radiosonne im Abstand *r* vom Scheibenzentrum in Sonnenradien; *rechte Skala* bezieht sich auf die Wellenlängen 1,4 und 3,7 m

(↑ 10.2.3.14.). Die im Koronaspektrum beobachteten Emissionslinien faßt man zur *L-Komponente* zusammen.

Korona und obere Chromosphäre sind der Entstehungsort der *solaren Röntgenstrahlung*. Auch die *Radiostrahlung* im Meterwellenbereich muß aus der Korona stammen, da diese Schicht für Strahlung dieser Wellenlänge undurchlässig ist. Die kürzerwellige Radiostrahlung wird jedoch in tieferen Schichten erzeugt. Der Durchmesser der „Radiosonne" ist daher wellenlängenabhängig (Abb. 10.2.2-3).

10.2.2.3. Sonnenaktivität. Alle veränderlichen, schnellebigen Phänomene auf der Sonne werden unter dem Begriff Sonnenaktivität zusammengefaßt. Die auffälligste und daher bekannteste Erscheinung dieser Art sind die *Sonnenflecken* (Tafel 54), Gebiete in der Photosphäre, die sich dunkel gegen die Umgebung abheben. Die größeren Sonnenflecken haben meist einen dunklen Kern, die *Umbra*, der von einem hellen Hof, der *Penumbra*, umgeben ist, die eine radiale Struktur aufweist. Der Durchmesserbereich der Sonnenflecken erstreckt sich von etwa 1000 km bei den gerade noch erkennbaren Poren bis zu rund 200000 km bei den größten Objekten. Neben Hauptflecken bilden sich häufig Nebenflecken aus; die Sonnenflecken treten meist in Gruppen auf. Der innerhalb einer Gruppe in der Richtung der Sonnenrotation vorangehende Hauptfleck wird *P-Fleck* genannt, während der im östlichen Teil der Gruppe be-

findliche nachfolgende Hauptfleck *F-Fleck* heißt. Die Lebensdauer eines Sonnenflecks liegt zwischen einem Tag und mehreren Monaten; dabei haben die kleineren Objekte durchschnittlich eine kürzere Lebenszeit. Die Häufigkeit der Sonnenflecken wird durch die *Fleckenrelativzahl* ausgedrückt, bei deren Berechnung sowohl die Anzahl der gerade sichtbaren Gruppen als auch die der einzelnen Flecken berücksichtigt wird. Es besteht ein deutlicher *Häufigkeitswechsel* mit einer mittleren Periode von etwa 11 Jahren (Abb. 10.2.2-4). Zu Beginn eines jeden Sonnenfleckenzyklus treten die Flecken hauptsächlich in heliographischen Breiten um $\pm 35°$ auf. Im Laufe der Zeit verlagert sich die Häufigkeitszone immer mehr zum Äquator, so daß die letzten Flecken eines Zyklus nur noch eine durchschnittliche heliographische Breite um $\pm 8°$ aufweisen.

Die Strahlungsintensität im Kern eines Flecks beträgt nur etwa 10% der Intensität einer gleich großen Fläche der ungestörten Photosphäre. Dieser Helligkeitsdifferenz entspricht eine wesentlich niedrigere mittlere Temperatur von 3300 K in den Flecken. In sie strömt von unten wirbelartig Materie ein, während nach oben Masse abgegeben wird. In den Flecken sind stets starke Magnetfelder vorhanden, deren Feldlinien aus den Fleckenkernen senkrecht, aus den Penumbren geneigt austreten. Die Feldstärke ist im Kern am größten und kann in der Größenordnung von 0,1 T liegen.

In den Fleckengruppen haben die beiden Hauptflecken in der Regel eine entgegengesetzte *Polarität*. Während eines Zyklus ist die Polarität der P-Flecken auf der nördlichen Sonnenhälfte gleichartig – es liegen z. B. nur Nordpole vor –, auf der südlichen Hemisphäre sind die P-Flekken dann auch untereinander gleichartig, aber entgegengesetzt zu denen der nördlichen Hemisphäre gepolt – es liegen dann z. B. nur Südpole vor. Beim nächsten Zyklus ist die Polung entgegengesetzt, so daß in Wirklichkeit kein 11jähriger, sondern ein 22jähriger *Zyklus* vorliegt.

Die Sonnenflecken sind nur eine Erscheinungsform der Sonnenaktivität. Sie treten immer in Verbindung mit den *Sonnenfackeln* auf, die aber auch ohne Flecken beobachtet werden, zum Teil wegen ihrer größeren mittleren Lebensdauer von mehreren Monaten.

Die Sonnenfackeln übertreffen ihre Umgebung an Helligkeit wegen einer Überhitzung hauptsächlich der Chromosphäre um mehrere 100 K gegenüber der Nachbarschaft. Die Fackeln bevorzugen die Fleckenzone. Gegen Ende eines Fleckenzyklus findet man kleine Fackeln aber auch in hohen heliographischen Breiten.

Mit besonderen Hilfsmitteln oder bei Sonnenfinsternissen sind an den Sonnenrändern helle Erhebungen zu beobachten, die *Protuberanzen* (Tafel 54). Diese Materiewolken sind Erscheinungen der Sonnenaktivität und ragen weit über die Chromosphäre hinaus. Vor der Sonnenscheibe sind sie nur in Spektroheliogrammen als dunklere längliche Filamente zu erkennen. Man findet sie besonders in den gleichen Zonen, in denen die Flecken und die Fackeln auftreten. Zu Beginn eines Zyklus beobachtet man die Protuberanzen vorwiegend bei $\pm 50°$ heliographischer Breite. Im Verlauf des Zyklus verlagern sich die Häufigkeitsmaxima immer mehr zum Äquator. Von diesen Hauptgebieten der Protuberanzentätigkeit spalten sich Nebenzonen ab, die in Richtung Pol wandern und mit dem Aktivitätsminimum absterben. Im ruhigen Zustand können die Protuberanzen mehrere Monate überdauern. Sie haben dann meist eine lamellenartige, langgestreckte Form und erheben sich wie Brückenbögen im Mittel 50000 km über die Chromosphäre; ihre Länge

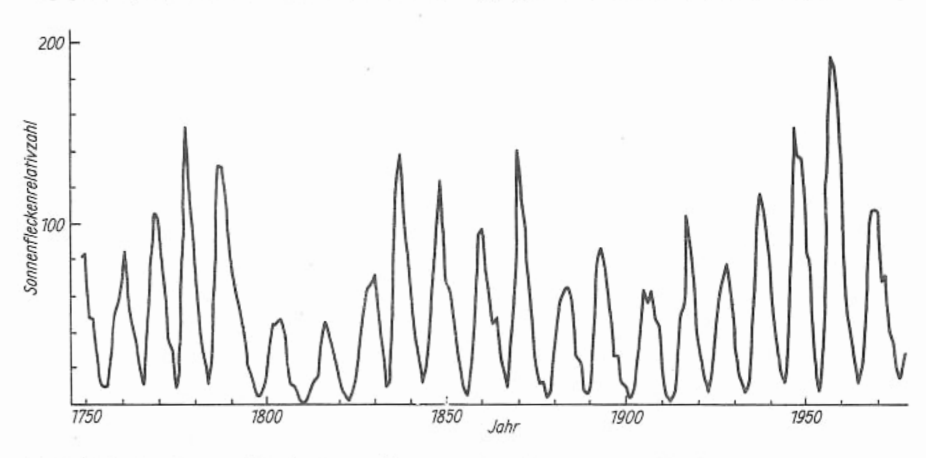


Abb. 10.2.2-4 Jahresmittel der Sonnenfleckenrelativzahlen von 1750 bis 1978

beträgt dabei über 150000 km und ihre Dicke rund 10000 km. In den Protuberanzen gibt es stets Strömungen, so daß ein ständiger Massenaustausch mit der Chromosphäre erfolgt. In den plötzlich einsetzenden Aktivitätsstadien verändern die Protuberanzen innerhalb weniger Stunden ihre Gestalt oder verschwinden sogar. Dabei wird oft seitlich auf gekrümmten Bahnen Masse an Gebiete der Chromosphäre abgegeben. Durch Massenzufluß auf den gleichen Bahnen erneuert sich des öfteren eine Protuberanz. Eine Aktivitätsphase ist oft mit einem *Protuberanzenaufstieg* verknüpft, bei dem Höhen von $2 \cdot 10^6$ km und Geschwindigkeiten von mehr als 500 $\mathrm{kms^{-1}}$ erreicht werden, die sich sprunghaft ändern können.

Über den Sonnenflecken können sich meist kurzlebige Protuberanzen als helle Knoten bilden, deren Masse in die Chromosphäre abfließt. Bei den *eruptiven Protuberanzen*, einer anderen Erscheinungsform, wird chromosphärisches Gas in die Korona geschleudert und dort zerstreut (Tafel 54).

Die *Sonneneruptionen* sind kurzzeitige Helligkeitsanstiege in begrenzten Gebieten der Photo- und Chromosphäre. Dem plötzlichen Intensitätszuwachs innerhalb weniger Sekunden oder Minuten folgt eine langsamere Helligkeitsabnahme.

Die Lebensdauer einer Eruption schwankt nach der Größe des Gebiets zwischen 10 und 90 min. Die größten Eruptionen können mehrere Quadratgrad der Sonnenoberfläche erfassen. Am Sonnenrand machen sich die Eruptionen als Ausweitungen der Chromosphäre in die Korona bemerkbar. Die größeren Eruptionen werden oft von eruptiven Protuberanzen begleitet. Die Häufigkeit der Eruptionen zeigt die gleiche Zeitabhängigkeit wie die anderen Erscheinungen der Sonnenaktivität.

In der Korona äußert sich die Sonnenaktivität in *Kondensationen* überhitzter Materie, die sich über mehrere Sonnenrotationen hinweg in Höhen von einigen 10000 km über den Sonnenflecken befinden. Auch ändert sich die Gesamtstruktur der Korona im Laufe des Aktivitätszyklus; sie weist zur Zeit des Sonnenfleckenmaximums lange Strahlen nach allen Richtungen auf und ist zur Zeit des Minimums an den Polen abgeflacht. Nur in Äquatornähe haben die Strahlen annähernd die gleiche Länge wie im Maximum.

Der *Radiofrequenzstrahlung*, die von der „ruhigen" Sonne ausgesendet wird, überlagert sich eine *Störstrahlung*, die eng mit den Aktivitätserscheinungen verknüpft ist. Die langsam veränderliche Komponente stammt aus den überhitzten Koronakondensationen. Sie wird im Meter- und Dezimeterwellenbereich beobachtet. Ihrer spektralen Energieverteilung nach ist sie thermischer Natur. Noch nicht mit optischen Erscheinungen identifizierte Bereiche sind die *R-Gebiete*, die sich 0,3 bis 1 Sonnenradius über der Photosphäre in der Korona befinden. Da die von diesen Gebieten ausgehende zirkular polarisierte Meterwellenstrahlung stark gebündelt ist, kann sie nur empfangen werden, wenn das betreffende R-Gebiet bei der Sonnenrotation in die Mitte der Sonnenscheibe gelangt. Die kurzzeitigen Störstrahlungen zeigen eine große Vielfalt. Zwar ist bekannt, daß sie nichtthermischen Ursprungs sind, doch ist noch nicht entschieden, ob sie durch Plasmaschwingungen erzeugt werden oder als Synchrotronstrahlung entstehen. Phänomenologisch teilt man die Strahlungen, deren Intensitäten gelegentlich um mehrere Zehnerpotenzen ansteigen, nach ihrem Spektrum in mehrere Klassen ein.

Zum *Typ I* werden Strahlungsstöße mit sehr schmalem Spektrum und sehr kurzer Dauer von 0,1 bis 0,5 s gerechnet. Die Ausbrüche vom *Typ II* treten häufig im Meter- und Dezimeterwellenbereich zu Beginn von Sonneneruptionen auf. Während eines Ausbruchs, der in der Regel einige Minuten anhält, verschiebt sich die Maximalintensität immer mehr zu größeren Wellenlängen. Wahrscheinlich wird diese Strahlung durch Plasmaschwingungen verursacht. In größerer Höhe werden wegen der geringeren Dichte die Plasmafrequenzen kleiner, wodurch die Wellenlängenverschiebung der Radiostrahlung erklärt werden kann. Aus der Zeitabhängigkeit der Wellenlängenverlagerung ergeben sich für die Geschwindigkeit, mit der sich die Strahlungsquelle von der Photosphäre entfernt, Werte von einigen 100 bis mehreren 1000 $\mathrm{kms^{-1}}$. Beim *Typ III* werden ähnliche Verschiebungen im Spektrum beobachtet wie bei den Ausbrüchen vom Typ II. Da diese Erscheinungen jedoch nur wenige Sekunden andauern, erhält man für die Geschwindigkeit der Strahlungsquelle 100000 $\mathrm{kms^{-1}}$, wie auf interferometrischem Wege direkt nachgewiesen wurde. Allerdings wurden nicht nur rasche Aufstiege in die Korona, sondern auch ähnlich schnelle Abwärtsbewegungen festgestellt. Als *Typ IV* werden Strahlungsausbrüche bezeichnet, bei denen die Intensitätssteigerung über einen weiten Bereich des Spektrums erfolgt. Diese Stöße haben eine Lebensdauer von mehreren Minuten bis zu einigen Stunden. Wie beim Typ II bewegen sich die Strahlungsquellen, ohne die Wellenlängen der Maximalintensität zu ändern, mit Geschwindigkeiten von einigen 100 $\mathrm{kms^{-1}}$ in die Korona, wo sie zum Stillstand kommen. Diese Ausbrüche werden meist einige Minuten nach Strahlungsstößen vom Typ II gegen Ende einer Eruption beobachtet. Die Strahlungsausbrüche vom *Typ V* treten nach solchen vom Typ III auf. Sie halten mehrere Minuten an. In ihren Eigenschaften sind sie den Strahlungsstößen vom Typ IV ähnlich.

Die Sonne sendet nicht nur elektromagnetische Strahlung der verschiedenen Wellenlängen, sondern auch *Teilchenstrahlung* aus, die vorwiegend aus Protonen, Heliumkernen und Elektronen besteht. Anders als das Plasma des Sonnenwinds bewegen sich diese scharf gebündelten *Teilchenströme* mit Geschwindigkeiten von einigen 1000 kms^{-1}. Wenn sie die Erde treffen, lösen sie *Ionosphärenstörungen, geomagnetische Stürme* und *Polarlichter* aus. Aus dem wiederholten Auftreten z. B. von geomagnetischen Stürmen in Abständen von 27 Tagen, der synodischen Rotationsdauer der Sonne, ergibt sich, daß ein Teilchenstrom mehrere Monate bestehen bleiben kann. Die Ursprungsgebiete solcher Korpuskularstrahlung, die man als *M-Gebiete* bezeichnet, konnten noch nicht eindeutig ermittelt werden.

10.2.3. Planetensystem

10.2.3.1. Planetenbewegung. Die neun großen Planeten Merkur, Venus, Erde, Mars, Jupiter, Saturn, Uranus, Neptun und Pluto bewegen sich im gleichen Sinn um die Sonne wie die Erde um ihre Achse rotiert, von West nach Ost. Mit Ausnahme des innersten und des äußersten Planeten, Merkur und Pluto, sind die Bahnebenen der Planeten nur schwach gegen die Ebene der Erdbahn geneigt, so daß das Planetensystem sehr flach ist. Alle Planetenbahnen sind *Ellipsen* mit der Sonne in einem der Brennpunkte. Außer bei Merkur und Pluto, weichen diese Bahnellipsen nur sehr wenig von Kreisbahnen ab. Die Planetenbewegung um die Sonne erfolgt so, daß der Leitstrahl Sonne – Planet in gleichen Zeitintervallen gleiche Flächen überstreicht, d. h., die Planeten bewegen sich in Sonnennähe, in der Nähe des Perihels ihrer Bahn, schneller als in Sonnenferne, in der Nähe des Aphels ihrer Bahn. Die Quadrate der Umlaufzeiten zweier Planeten verhalten sich zueinander wie die Kuben der großen Halbachsen der Bahnen. Die genannten Eigentümlichkeiten der Planetenbewegung sind Inhalt der drei *Keplerschen Gesetze* und resultieren letztlich aus dem *Newtonschen Gravitationsgesetz.*

Mit Ausnahme von Neptun und Pluto befolgen die Planeten eine regelmäßige Abstandsreihe, die *Titius-Bodesche Reihe*, die auch als Formel angegeben werden kann. Sie läßt allerdings zwischen Mars und Jupiter eine Lücke in der Planetenfolge erkennen. Statt eines großen Planeten bei 2,8 AE beobachtet man zahlreiche kleine Planeten, auch *Planetoiden* oder *Asteroiden* genannt, mit großen Bahnhalbachsen zwischen 2,17 und 3,3 AE.

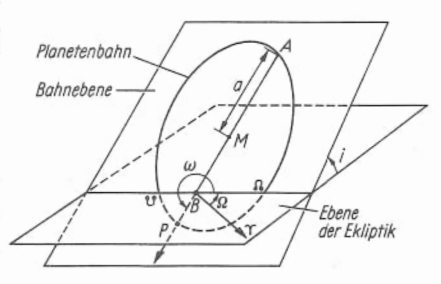


Abb. 10.2.3-2 Bahnelemente der Planetenbahn; *M* Mittelpunkt, *B* Brennpunkt der Bahnellipse, ☊ aufsteigender, ☋ absteigender Knoten, ♈ Frühlingspunkt, *P* Perihel, *A* Aphel, *a* große Halbachse, *i* Neigung der Bahnebene, Ω Länge des aufsteigenden Knotens, ω Abstand des Perihels vom aufsteigenden Knoten; neben i, a, Ω und ω gehören noch die numerische Exzentrizität e ($e = MB/a$) und eine Durchgangszeit durch das Perihel zu den Bahnelementen

Zwischen den Bahnen einiger benachbarter Planeten bestehen *Resonanzbeziehungen*, d. h., ihre Umlaufzeiten stehen im Verhältnis kleiner ganzer Zahlen zueinander, z. B. $U_{♃} : U_{♄} = 2 : 5$ (Tab. 10.2.3-1, Abb. 10.2.3-2).

Tab. 10.2.3-1 Bahndaten der Planeten

Planet	Symbol	Große Halbachse in AE	Umlaufzeit in a	Numerische Exzentrizität	Neigung der Bahn zur Ekliptik in Grad
Merkur	☿	0,387	0,2408	0,2056	7,0
Venus	♀	0,723	0,6152	0,0068	3,4
Erde	♁	1,000	1,0000	0,0167	0,0
Mars	♂	1,524	1,881	0,0934	1,8
Jupiter	♃	5,203	11,86	0,0485	1,3
Saturn	♄	9,54	29,46	0,0556	2,5
Uranus	♅	19,18	84,0	0,0472	0,8
Neptun	♆	30,06	164,8	0,0086	1,8
Pluto	♇	39,7	248,4	0,253	17,1

10.2.3.2. Beschaffenheit der Planeten. Nach ihren mittleren Dichten kann man zwei Gruppen von Planeten definieren, die sich markant durch ihre chemische Zusammensetzung unterscheiden und nach ihren Prototypen *terrestrische* oder *erdartige* und *jupiterartige* Planeten genannt werden. Die terrestrischen Planeten – Merkur, Venus, Erde und Mars – weisen relativ kleine Massen, kleine Radien, aber hohe mittlere Dichten auf (↑ Abb. 10.2.3-11). Wie die Modellrechnungen zeigen, besitzen sie einen ähnlichen *Schalenaufbau* wie die Erde. Ein *Kern* aus metallischem Eisen, beim Mars möglicherweise aus Eisensulfid, das sich zum Teil in flüssigem Zustand befindet, wird von einem *Mantel* aus schweren Oxiden und basischen Silicaten umgeben, den eine *Kruste* leichteren Silicatgesteins nach außen abschließt (Abb. 10.2.3-3). Kennzeichnend für die terrestrischen Planeten sind relativ *dünne Atmosphären* bzw. Ionosphären und mit Ausnahme der Erde *schwache Magnetfelder*. Zu diesen Planeten gehören insgesamt nur drei Monde.

Bei den wesentlich größeren und massereicheren jupiterartigen Planeten muß man die beiden *Riesenplaneten* Jupiter und Saturn gesondert betrachten. Sie sind im Chemismus sonnenähnlich, d. h. überwiegend aus Wasserstoff und Helium zusammengesetzt. Zumindest beim Jupiter geht der flüssige Wasserstoff, sobald der Druck ungefähr 300 GPa (3 Mbar) erreicht, in eine metallische Modifikation über. Wahrscheinlich existiert im Jupiterzentrum noch ein fester Kern von erdartiger Zusammensetzung, über dessen physikalischen Zustand in Anbetracht des extremen Drucks von 10 TPa (100 Mbar) nichts bekannt ist.

Uranus und Neptun, die masse- und größenmäßig eine Zwischenstellung zwischen den terrestrischen und den Riesenplaneten einnehmen, scheinen neben normalem Wasserstoff und Helium auch größere Mengen der nächst schwereren kosmisch häufigen Elemente, z. B. Kohlenstoff, Sauerstoff, Stickstoff und Neon, in sich zu vereinigen. Möglicherweise besitzen auch sie erdartige Kerne. Über den Aufbau des Plutos gibt es noch keine diskutablen Modellvorstellungen, weil Masse und Radius dieses Planeten noch nicht sicher bekannt sind.

Charakteristisch für die jupiterartigen Planeten sind Rotationsdauern von weniger als 1 d, dichte, ausgedehnte und wolkenreiche *Atmosphären*, deren Hauptkomponenten H_2, He und Wasserstoffverbindungen sind, sowie *Ringe* und *Mondsysteme*. Jupiter und Saturn besitzen starke und ausgedehnte *Magnetosphären*. Jupiter, Saturn und Neptun verfügen über *innere Energiequellen*, denn sie strahlen im Infraroten mehr Energie ab, als sie von der Sonne in Form von Licht empfangen (Tab. 10.2.3-4).

10.2.3.3. Merkur. Der sonnennächste Planet Merkur benötigt für eine Umdrehung um seine Achse genau 2/3 der Umlaufzeit um die Sonne. Diese Resonanz zwischen Rotation und Bahnbewegung hat zur Folge, daß sich jeder Punkt der Merkuroberfläche 88 d auf der Sonnenseite und 88 d auf der Nachtseite befindet.

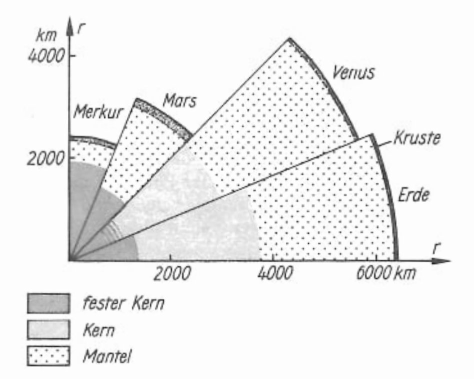


Abb. 10.2.3-3 Innerer Aufbau der erdartigen Planeten; die Krustendicke ist nur bei der Erde sicher bekannt

Tab. 10.2.3-4 Physikalische Eigenschaften der Planeten, * unsichere Werte

Planet	Äquator-radius in km	Masse in Mt	Mittlere Dichte in kg m^{-3}	Abplattung	Rotations-periode in d	Magnetisches Moment in T m^3
Merkur	2439	0,0553	5440	0	58,6	$4{,}9 \cdot 10^{12}$
Venus	6052	0,8150	5240	0	−243,0	$1 \cdot 10^{12}$
Erde	6378	1,0000	5520	0,0034	1,00	$8{,}0 \cdot 10^{15}$
Mars	3397	0,1074	3960	0,0132	1,02	$2 \cdot 10^{12}$*
Jupiter	71398	317,89	1320	0,061	0,41	$1{,}4 \cdot 10^{20}$
Saturn	60000	95,17	690	0,096	0,43	$1 \cdot 10^{20}$
Uranus	25400	14,56	1340	0,06	1,0*	
Neptun	24300	17,24	1750	0,02	0,9*	
Pluto	3000*	0,0017*	700*	0*	6,39	

Die Zeitspanne zwischen zwei Sonnenaufgängen, der *Merkurtag*, ist also genau doppelt so lang wie das *Merkurjahr*. Der Merkur empfängt pro Flächeneinheit im Mittel 6mal, im Perihel sogar 10mal soviel Sonnenlicht wie die Erde. An der Oberfläche des praktisch atmosphärelosen Himmelskörpers herrschen *Mittagstemperaturen* zwischen 570 und 700 K.
Wie Modellrechnungen zeigen, muß der Merkur einen relativ großen *Eisenkern* besitzen, der 60 bis 70% der Merkurmasse enthält. Von ihm geht das *Magnetfeld* dieses Planeten aus, dessen Feldstärke an der Oberfläche rund 1% des irdischen Werts beträgt. Der *Gesteinsmantel* des Merkurs ist demnach nur etwa 700 km dick (↑ Abb. 10.2.3-3).
Dominierendes Merkmal der *Merkuroberfläche* sind wie beim Mond die Einschlagskrater. Auch mareartige Flächen, sogar Zirkularmaria, sind vorhanden. Es gibt geologische Anzeichen dafür, daß der Planet in seiner Geschichte einen *Schrumpfungsprozeß* durchgemacht hat. Heute ist der Merkur wie der Mond ein geologisch toter Himmelskörper (Tafel 55).
10.2.3.4. Venus. Nach Größe und Masse ist die Venus der erdähnlichste Planet. Da die dichte Wolkendecke den Blick auf die Oberfläche verhindert, blühten in der Vergangenheit die Spekulationen über die dort herrschenden Bedingungen. Auch gelang es erst 1965 auf radarastronomischem Wege die *Rotationsdauer* der Venus zu bestimmen. Der Planet dreht sich in 243 d einmal um seine Achse, und zwar *entgegen* dem im Sonnensystem üblichen Drehsinn.
Planetensonden ermittelten an den verschiedenen Landeorten für den Venusboden Temperaturen zwischen 720 und 750 K, wobei keine Schwankungen zwischen Tag und Nacht auftreten, und einen atmosphärischen Druck von 8,6 bis 9,5 MPa (86 bis 95 bar). Die untere Atmosphäre ist ziemlich staubfrei, extrem trocken und besteht zu 96 Vol.-% aus CO_2 (Tab. 10.2.3-5). Diese Zusammensetzung und die dichte Wolkendecke bewirken einen starken *Treibhauseffekt*, weil die vom Venusboden abgegebene Wärmestrahlung von der Atmosphäre absorbiert wird, so daß sich Boden und untere Atmosphäre stark aufheizen müssen. Während am Venusboden fast Windstille herrscht, tobt in Wolkenhöhe orkanartiger Sturm, durch den die Wolkendecke den Planeten in 4 d einmal umrundet. Die *Venuswolken* befinden sich im Höhenbereich von 35 bis 68 km und bestehen zumindest in ihren dichtesten Schichten (48 bis 52 km) aus Schwefelsäuretröpfchen.
Die Venus besitzt eine *Ionosphäre* mit einer maximalen Elektronendichte von $5{,}6 \cdot 10^{12}\,m^{-3}$ am Tage und $2 \cdot 10^{10}\,m^{-3}$ in der Nacht in 142 km Höhe. In etwa 500 km Höhe geht die Elektronendichte stark zurück (Ionopause), wahrscheinlich durch die Wechselwirkung zwischen Ionosphäre und Sonnenwind (↑ 10.2.3.14.). Wie die Erde ist auch die Venus von einer *Wasserstoffkorona* umgeben.
Von der Radarastronomie wurden auf der Venus Krater, gebirgige Flächen, ein Riesenvulkan, ein 1500 km langer Graben und andere Strukturen entdeckt. Aufnahmen der Sonde Venus 9 zeigen zahlreiche größere Steine und Blöcke am Venusboden, während im Landegebiet von Venus 10 das anstehende Gestein zu sehen ist (Tafel 55). Die Venus sollte im Inneren ähnlich aufgebaut sein wie die Erde (↑ Abb. 10.2.3-3). Daß sie kein nachweisbares Magnetfeld besitzt, hängt wahrscheinlich mit der langsamen Rotation zusammen.
10.2.3.5. Mars. An der Oberfläche des Planeten Mars beobachtete man seit dem 17. Jh. Phänomene, die Parallelen zu irdischen Gegebenheiten darstellten oder sich als solche deuten ließen: die Tageslänge, die Jahreszeiten, die weißen Polkappen, die rötlichen Wüstengebiete mit ihren Sandstürmen, die jahreszeitlichen Variationen der dunklen Gebiete, die auf Vegetation hinzuweisen schienen, die „Kanäle", die spekulativ als künstliche Wasserstraßen gedeutet wurden, und die winzigen Monde, die sogar zeitweise

Tab. 10.2.3-5 Chemische Zusammensetzung der Troposphäre der erdartigen Planeten (Angaben in Volumen-%), * variabel, teilweise als Kondensat

Komponente	Venus	Erde	Urerde (nach Rubey)	Mars
CO_2	97	0,03	91	95
N_2	2	78,09	6,4	2,7
O_2	0,1	20,95	0	0,15
Ar		0,93		1,6
H_2O	1	1–2*		
sonstige	CO	Ne	H_2S	CO
	HCl	He	HCl	O_3
	HF	Kr		Kr
		Xe		Xe
Wolkenmaterial	H_2SO_4	H_2O	H_2O	H_2O, CO_2

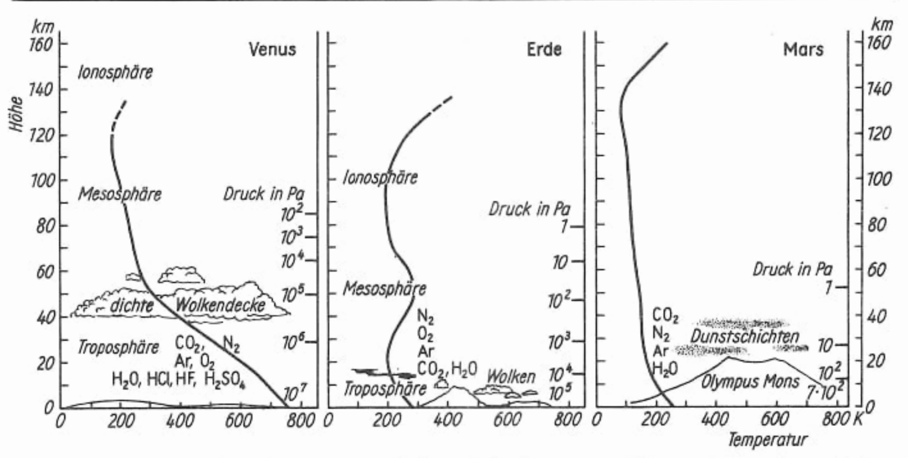


Abb. 10.2.3-7 Verlauf von Temperatur und Druck in den Atmosphärenschichten der erdartigen Planeten; die Druckwerte beziehen sich auf die Höhen, in denen sie eingezeichnet sind, der unterste auf die Planetenoberfläche

als künstliche Himmelskörper angesehen wurden.

Die von zahlreichen Planetensonden gewonnenen Ergebnisse haben inzwischen das spekulative Marsbild korrigiert. Der Planet nimmt danach eine Zwischenstellung zwischen Mond und Erde ein. Auf ihm hat es nie höher entwickelte Lebensformen gegeben.

Die *Marsoberfläche* zeigt kontinentale Flächen mit zahlreichen *Einschlagskratern* einschließlich Zirkularmaria und mareartige Flächen, die aber im Unterschied zu den Maria des Monds meist hell sind und zahlreiche *Vulkanbauten* tragen. Auf dem Mars findet man die größten Schildvulkane des Sonnensystems, z. B. Olympus Mons mit einem Basisdurchmesser von 600 km und einer Höhe relativ zur Umgebung von etwa 25 km. Um derartige Massive tragen zu können, muß die Marskruste wesentlich dicker und starrer als die Erdkruste sein. Das Fehlen direkter Anzeichen für bewegliche Platten, wie sie für die Lithosphäre der Erde charakteristisch sind, unterstreicht diese Folgerung. Die auffälligste tektonische Erscheinung ist ein sich etwa 4000 km in ost-westlicher Richtung erstreckendes System von 100 bis 200 km breiten und stellenweise bis zu 6 km tiefen *Cañons*, das den Namen Valles Marineris erhielt. Geologische Strukturen, die dem Phänomen der *Marskanäle* zugrunde liegen könnten, wurden nicht entdeckt. Die von manchen Astronomen beobachteten dunklen Linien an der Wahrnehmbarkeitsgrenze beruhen offensichtlich auf Täuschungen (Tafel 59).

Auf dem Mars wurden zahlreiche *Flußbetten*, breite *Stromtäler* und andere auf fließendes Wasser zurückgehende *Erosionserscheinungen* entdeckt. Die Zeiträume, in denen auf dem Mars Wasser in flüssiger Form einen wesentlichen geologischen Faktor darstellte, können jedoch nur relativ kurz gewesen sein, denn man fand in den Niederungen der Marskruste keine dicken Sedimentschichten. Eine mit der Erde vergleichbare Hydrosphäre ist offenbar nie vorhanden gewesen. Heute ist es auf dem Mars extrem trocken, der Wasserdampfgehalt der Atmosphäre beträgt nur 10 bis 80 µm Niederschlagswasser. In den Zentralgebieten der Polkappen kommt jedoch neben dem dominierenden CO_2-Schnee auch H_2O-Schnee bzw. -Eis vor. Es gibt sogar Anzeichen für eine zeitweise Vergletscherung der Polargebiete. Ebenso finden sich Anzeichen für die Existenz gefrorenen Wassers im Marsboden. Da die mittlere Temperatur auf dem Mars bei etwa 230 K liegt, herrscht im Marsboden *Dauerfrost*.

Tab. 10.2.3-6 Chemische Zusammensetzung des Marsgesteins nach Keil und Mitarbeitern in Masse-%

	Chryse Planitia	Utopia Planitia	Fehlergrenzen
SiO_2	44,7	42,8	± 5,3
Al_2O_3	5,7		1,7
Fe_2O_3	18,2	20,3	2,9
MgO	8,3		4,1
CaO	5,6	5,0	1,1
TiO_2	0,9	1,0	0,3
SO_3	7,7	6,5	7,3
Cl	0,7	0,6	0,3

Starke Trübungen der Marsatmosphäre waren bereits früher als Zeichen von *Windtätigkeit* erkannt worden. Auf den Nahaufnahmen findet man Dünenfelder, vom Wind abgelagerte, geschichtete Sedimente und durch Windverwehung feinen Bodenmaterials bewirkte Albedoveränderungen im Laufe weniger Tage. Es kann als sicher gelten, daß die beobachteten jahreszeitlichen Variationen der dunklen Gebiete ein Windeffekt sind.

Wie der Boden des Monds hat auch der des Mars Regolithcharakter. Die ziegelrote Färbung des Marsbodens geht wahrscheinlich auf den hohen Gehalt an Eisenoxid, vor allem Hämatit, zurück. In den Landegebieten der Vikingsonden, in Chryse Planitia und Utopia Planitia, ist der Boden von zahlreichen Steinen bedeckt, unter denen viele das Aussehen von Windkantern haben, wie sie in irdischen Sandwüsten vorkommen. Organisches Material oder Mikroorganismen wurden im Boden nicht gefunden. Bei den biologischen Experimenten stellte sich jedoch eine unerwartet hohe chemische Aktivität des Marsbodens heraus, die auf einen sehr hohen Oxydationsgrad schließen läßt (Tab. 10.2.3-6).

In den genannten Landegebieten beträgt der Luftdruck rund 700 Pa (7 mbar). In Chryse Planitia erreicht die Temperatur kurz vor Sonnenaufgang ein Minimum von 187 K, am frühen Nachmittag ein Maximum von 242 K. Die Windgeschwindigkeit beträgt im Mittel 2,4 ms^{-1}. Die Marsatmosphäre besitzt einen hohen Staubgehalt. Anders als bei der Erde ist auf dem Mars die Farbe des Tageshimmels rosa, ein Umstand, der durch die Streuung des Lichts an den feinen Aerosolteilchen zustande kommt (Abb. 10.2.3-7).

Der Mars besitzt eine *Ionosphäre* mit einer maximalen Elektronendichte von 10^{11} m^{-3} auf der Tagseite in einer Höhe zwischen 123 und

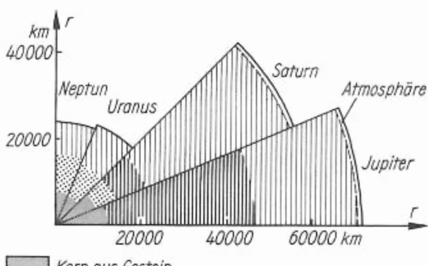


Abb. 10.2.3-8 Innerer Aufbau der jupiterartigen Planeten; *r* Abstand vom Zentrum; die Dicken der Atmosphären sind unbekannt

155 km. Wie Venus und Erde umgibt auch ihn eine *Wasserstoffkorona*. Der Planet besitzt ein schwaches *Magnetfeld*, das auf den kleinen Kern zurückzuführen ist, der möglicherweise aus Eisen und Eisensulfid besteht (↑ Abb. 10.2.3-4).

Die beiden *Marssatelliten* Phobos und Deimos sind eindeutig natürliche Himmelskörper (↑ 10.2.3.11., Tafel 59).

10.2.3.6. Jupiter. Der infolge seiner schnellen Rotation auffällig abgeplattete Riesenplanet Jupiter besitzt neben planetaren auch stellare Eigenschaften, z. B. die sonnenähnliche chemische Zusammensetzung, die Energieproduktion im Inneren (s. a. 10.2.3.2.) und die differentielle Rotation, durch die die äquatornahen Gebiete schneller rotieren als die in höheren Breiten gelegenen (s. a. 10.2.2.1.). Die Masse des Planeten ist jedoch um rund einen Faktor 80 zu klein, um im Zentrum die für Sterne typischen Kernfusionen zu zünden.

Nach den verschiedenen Modellen über das Jupiterinnere besitzt der Planet eine Zentraltemperatur zwischen 13000 und 35000 K. Diese Temperatur ist hoch genug, daß der Wasserstoff, auch der metallische (↑ 10.2.3.2.), flüssig ist. Als feste Oberfläche läßt sich höchstens die Peripherie des festen Kerns aus schweren Elementen auffassen, dessen Radius wahrscheinlich bei weniger als 20% des Planetenradius liegt und in dem 10 bis 20 Erdmassen vereinigt sein dürften (Abb. 10.2.3-8).

In den von der Erde aus sichtbaren Wolkenschichten der ausgedehnten Atmosphäre, in der neben H_2 und He Ammoniak (NH_3), Methan (CH_4), Ethan (C_2H_6), Ethin (C_2H_2), Phosphin (PH_3) u. a. Wasserstoffverbindungen entdeckt wurden, beträgt der Gasdruck ungefähr 70 kPa (700 mbar) und die Temperatur liegt bei 150 K. Die Temperatur ist auf der Tag- und Nachtseite gleich groß. Die Wolken bestehen wahrscheinlich aus Kriställchen bzw. Tröpfchen aus NH_3 und Ammoniumhydrogensulfid (NH_4SH). Die Jupiterionosphäre ist etwa 1500 km dick.

Durch die schnelle Rotation bilden sich die für den Jupiter typischen äquatorparallelen Wolkenstreifen aus. Die hellen Zonen liegen etwa 20 km höher und sind im Mittel um 9 K kühler als die dunklen Bänder oder Gürtel. Breite, Begrenzung und Farbton der Zonen und Bänder sind zeitlich variabel. Wie Raumsondenaufnahmen des Planeten zeigen, hört die Streifenstruktur in hohen Breiten auf.

Das markanteste Detail auf der sichtbaren Jupiterscheibe ist der *Große Rote Fleck* in der südlichen tropischen Zone (Tafel 56). Dieser bedeckt eine Fläche von der Größe der Erdoberfläche und driftet langsam in Länge. Seine Färbung kann zeitweise so stark ausbleiben, daß er unsichtbar wird. Wahrscheinlich liegt

diesem Fleck ein langlebiger Wirbel, ein *Zyklon*, in der Atmosphäre zugrunde.
Der Jupiter besitzt die stärkste *Magnetosphäre* des Planetensystems. Auf der Sonnenseite reicht sie bis zu rund 100 Jupiterradien Abstand, der *Magnetschweif* auf der sonnenabgewandten Seite konnte noch jenseits der Saturnbahn nachgewiesen werden. Das Magnetfeld ist entgegengesetzt gepolt und an der Planetenoberfläche rund 30mal so stark wie das der Erde. Nur in der Nähe des Planeten ähnelt es einem Dipolfeld, in dem sich rund 1000mal intensivere *Strahlungsgürtel* als in der irdischen Magnetosphäre befinden. Das magnetische Moment des Jupiters beträgt $1{,}4 \cdot 10^{20}$ Tm³. Aufrechterhalten wird dieses Magnetfeld durch

Tab. 10.2.3-9 Satelliten des Sonnensystems; [1] unsichere Werte, [2] Bahnneigung des Erdmonds relativ zur Ekliptik, bei Himalia, Elara, Pasiphae, Sinope, Lysithea, Carme, Ananke, Leda, Phoebe und Nereide relativ zur Bahnebene des Planeten und bei allen anderen relativ zur Äquatorebene des Planeten, [3] Massen von Phobos und Deimos betragen $1{,}0 \cdot 10^{16}$ bzw. $2 \cdot 10^{15}$ kg, [4] Existenz von Janos unsicher

Satellit	Große Bahnhalbachse in 10^3 km	Numerische Exzentrizität	Bahnneigung in °	Radius in km	Masse in 10^{20} kg
Erde:					
Erdmond	384,4	0,055	5,1[2]	1738	735
Mars:					
Phobos	9,4	0,015	1,1	12	[3]
Deimos	23,5	0,001	0,9–2,7	6	[3]
Jupiter:					
Amalthea	181,3	0,003	0,4	120	
Io	421,6	0,000	0,0	1820	891
Europa	670,9	0,000	0,5	1500	487
Ganymed	1070	0,001	0,2	2635	1490
Kallisto	1880	0,01	0,2	2500	1065
Leda	11110	0,146	26,7	5[1]	
Himalia	11470	0,158	27,6	85	
Lysithea	11710	0,130	29,0	12[1]	
Elara	11740	0,207	24,8	40	
Ananke	20700	0,17	147	10[1]	
Carme	22350	0,21	164	15[1]	
Pasiphae	23300	0,38	145	18[1]	
Sinope	23700	0,28	153	14[1]	
Saturn:					
Janus[4]	159,6	0	0	110[1]	
Mimas	186	0,020	1,5	170[1]	0,37
Enceladus	238	0,004	0,0	250[1]	0,85
Tethys	295	0,000	1,1	500[1]	6,26
Dione	377	0,003	0,0	575	11,6
Rhea	527	0,001	0,4	800	18,2
Titan	1222	0,024	0,3	2900	1401
Hyperion	1481	0,104	0,4	110[1]	
Iapetus	3560	0,028	14,7	800	22,4
Phoebe	12930	0,163	150	40[1]	
Uranus:					
Miranda	130	0,017	3,4	150[1]	
Ariel	192	0,003	0	400[1]	
Umbriel	267	0,004	0	275[1]	
Titania	438	0,002	0	500[1]	
Oberon	586	0,001	0	450[1]	
Neptun:					
Triton	354	0,00	160	1600[1]	3400[1]
Nereide	5510	0,75	27,6	150[1]	
Pluto					
Charon	20[1]	0[1]		600[1]	

Zirkulationsströmungen im metallischen Wasserstoff des Planeten. Die Jupitermagnetosphäre ist eine Quelle kosmischer Strahlung.

Die inneren *Jupitermonde* (↑ 10.2.3.11., Tafel 57) bewegen sich innerhalb der Magnetosphäre und beeinflussen sie, z. B. löst die Io an einer bestimmten Stelle ihrer Bahn Ausbrüche im Dekameterwellenbereich (Wellenlänge ≧ 10 m) aus. Von den 13 Jupitermonden, die Namen tragen (Tab. 10.2.3-9), sind die Bahnen sicher bekannt. Es wurden aber mindestens drei weitere Jupitersatelliten gesichtet. Innerhalb der Bahn der Amalthea wurde 1979 ein schmaler *Ring* entdeckt.

10.2.3.7. Saturn. Der Riesenplanet Saturn ähnelt in vielen Merkmalen dem Jupiter. In seiner *Atmosphäre* wurden neben Wasserstoff Methan, Ethan und Phosphin nachgewiesen. Der Methananteil ist größer als der in der Jupiteratmosphäre. Die *Wolkendecke* zeigt ein einfacheres System von Zonen und Bändern und blassere Farbtöne als die des Jupiters. Gelegentlich treten auf dem Saturn *weiße Flekken* mit relativ kurzer Lebenszeit auf. Wegen der niedrigeren Schwerebeschleunigung nimmt die Gasdichte in der Saturnatmosphäre mit der Höhe langsamer ab als in der Jupiteratmosphäre. Die *Temperatur* in der oberen Wolkendecke beträgt 97 K.

Der Saturn besitzt eine *Magnetosphäre*. Das Magnetfeld hat in guter Näherung Dipolcharakter, Rotations- und Magnetfeldachse stimmen (anders als bei der Erde und dem Jupiter) überein, die Polung ist analog der der Erde. Das magnetische Moment des Saturns ist allerdings 30mal geringer als das des Jupiters.

Das *Ringsystem* des Saturns besteht aus zahllosen Brocken, die den Planeten genau in seiner Äquatorebene umkreisen. Im Fernrohr sichtbare Lücken (*Teilungen*) zerteilen den Ring in konzentrische Teilringe. Mit Hilfe von Planetensonden wurden außerhalb des A-Rings weitere Ringe entdeckt, und die bisher bekannten hellen Ringe erwiesen sich als reichhaltig strukturiert. Der Saturn besitzt zehn *Monde* mit gut bekannten Bahnen (↑ Tab. 10.2.3-9). Auf den Sondenaufnahmen wurden jedoch weitere Monde gesichtet (↑ 10.2.3.11., Tafel 58).

10.2.3.8. Uranus. Durch die *Neigung der Rotationsachse* um 98° zur Bahnebene entstehen beim Uranus eine Reihe von planetarischen Kuriositäten. Beispielsweise ist es zu den Zeiten, in denen ein Pol zur Sonne zeigt, in dem betreffenden Polargebiet wärmer als am Uranusäquator. Dieser Umstand wird 1985 eintreten, wenn der Uranus seinen Nordpol der Sonne zukehrt. Wegen der großen Neigung der Achse unterliegt die Länge des Uranustags, definiert als Zeitspanne zwischen zwei Sonnenaufgängen, extremen Schwankungen im Laufe des Uranusjahrs.

Bisher konnte nicht eindeutig nachgewiesen werden, ob die Uranuswolkendecke auch Streifen zeigt. In der Atmosphäre des Planeten fand man Wasserstoff und Methan. Die Temperatur in Wolkenhöhe wurde zu 57 K bestimmt.

Der Uranus besitzt ein *Ringsystem*, das aus mindestens neun sehr schmalen Ringen besteht. Es wurde 1976 bei der Beobachtung einer Sternbedeckung durch den Uranus entdeckt. Der Uranus wird von fünf *Monden* umlaufen (↑ 10.2.3.11.).

10.2.3.9. Neptun. Der Neptun ist dem Uranus sehr ähnlich. Neben Wasserstoff und Methan wurde in seiner Atmosphäre auch Ethan nachgewiesen. Die Temperatur in Wolkenhöhe beträgt 55 K. Wegen seiner Kleinheit konnten bisher auf dem Neptunscheibchen keinerlei Strukturen festgestellt werden. Die mit spektroskopischen und fotometrischen Methoden abgeleitete Rotationsdauer des Planeten muß noch als sehr unsicher betrachtet werden. Auch beim Neptun gibt es erste Anzeichen für eine *innere Energiequelle* und für eine *Magnetosphäre*. Der Planet besitzt zwei *Monde* (↑ 10.2.3.11.).

10.2.3.10. Pluto. Masse und Radius des äußersten Planeten sind noch nicht genau bekannt, so daß der Wert der mittleren Dichte noch zu unsicher ist, um irgendwelche Aussagen über seine chemische Natur machen zu können. An der *Oberfläche* dieses wahrscheinlich atmosphärelosen Planeten, deren Temperatur bei 45 K liegt, wurde auf spektroskopischem Wege festes Methan nachgewiesen. Der Pluto besitzt einen *Mond* (↑ 10.2.3.11.).

10.2.3.11. Satelliten. Mit Ausnahme von Merkur und Venus besitzen alle Planeten natürliche Satelliten, die jupiterartigen Planeten sogar ausgeprägte *Satellitensysteme*. Bisher sind 34 derartige Himmelskörper katalogisiert (↑ Tab. 10.2.3-9), entdeckt wurden jedoch bisher mehr als 40. Die größten von ihnen, der Saturnmond Titan, der Jupitermond Ganymed und möglicherweise auch der Neptunmond Triton, sind größer als der Planet Merkur. Diese großen Monde besitzen auch Atmosphären. Die kleinsten Satelliten liegen mit Durchmessern von wenigen Kilometern nur knapp über der heutigen Nachweisgrenze. Streng genommen gehören auch die zentimeter- oder metergroßen Brocken der Ringe von Jupiter, Saturn und Uranus zu den Satelliten.

Von den 13 *Jupitermonden* der Tab. 10.2.3-9 bewegen sich die fünf innersten, Amalthea, Io, Europa, Ganymed und Kallisto, auf Kreisbahnen in der Äquatorebene. Sie rotieren gebunden wie der Erdmond und zeigen eine regelmäßige Abstandsfolge vom Jupiter wie die Planeten von der Sonne. Ihre mittlere Dichte nimmt mit zunehmendem Abstand vom Jupiter ab. Die restlichen acht Satelliten bilden nach

den Bahndaten zwei Gruppen. 1.: Leda, Himalia, Lysithea, Elara. 2.: Ananke, Carme, Pasiphae, Sinope. Die Körper der 2. Gruppe umlaufen den Jupiter entgegen dem üblichen Drehsinn, sie sind rückläufig. Bei den meisten dieser kleinen Körper dürfte es sich um *eingefangene Asteroiden* handeln.
Das ausgeprägteste Satellitensystem besitzt der Saturn. Lediglich Iapetus und Phoebe zeigen Abweichungen, ersterer durch seine Bahnneigung, letzterer durch seine Rückläufigkeit. Die Existenz von Ianos ist umstritten. Die *Saturnmonde* sind durchweg größere Körper. Titan ist der größte Satellit des Sonnensystems und besitzt eine dichtere Atmosphäre als der Planet Mars.
Ein sehr regelmäßiges System bilden auch die fünf *Monde des Uranus*. Die zwei *Neptunmonde* fallen durch extreme Bahneigenschaften auf. Während Triton eine Kreisbahn beschreibt, bewegt sich Nereide auf der am stärksten elliptischen Bahn eines Monds überhaupt (Abb. 10.2.3-10).

10.2.3.12. Kleine Planeten. Seit 1801 fand man mehrere Tausend kleine Planeten, auch *Planetoiden* oder *Asteroiden* genannt, die die „Planetenlücke" zwischen Mars und Jupiter bevölkern (s. a. 10.2.3.1.). Sie bilden dort den *Planetoiden- oder Asteroidengürtel*. Bei etwa 2200 von ihnen ist die Bahn sicher erfaßt. Diese Objekte tragen in der Reihenfolge ihrer Entdeckung eine Nummer und einen Namen, der in der Regel auf den Entdecker zurückgeht. Mit abnehmender Größe nimmt die Anzahl der Planetoiden stark zu. Aus der statistischen Auswertung fotografischer *Suchprogramme* schätzte man ab, daß es insgesamt 500000 Planetoiden gibt, die in der für ihre Entdeckung günstigsten Stellung, in der *Opposition*, heller als 21,2 Größenklassen sind. Die kleinsten bekannten Planetoiden sind unregelmäßig begrenzte Brokken von knapp Kilometergröße, die nur deswegen entdeckt wurden, weil sie der Erde besonders nahe kamen. Die Gesamtmasse aller Planetoiden beträgt nur $2{,}4 \cdot 10^{21}$ kg, d. h. nur 0,0004 Erdmassen.
Die Häufigkeitsverteilung der großen Bahnhalbachsen zeigt, daß es Abstände von der Sonne gibt, die von den Planetoiden gemieden werden, die sogenannten *Kirkwoodlücken* des Planetoidengürtels. Sie markieren jene Sonnenabstände, in denen Resonanz zur Jupiterbewegung besteht. Es gibt allerdings auch Resonanzstellen, in denen eine Häufung von Planetoiden auftritt. Im Planetoidengürtel finden sich Gruppen von Objekten, die durch fast identische Bahnelemente ihre Zusammengehörigkeit zeigen, die *Hirayama-Familien* oder *Jetstreams*.

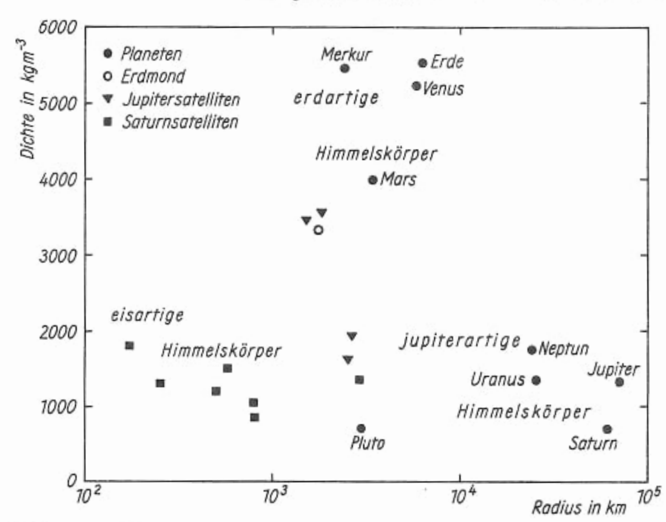


Abb. 10.2.3-10 Radius-Dichte-Diagramm der Planeten und einiger großer Satelliten

Neben den Planetoiden des Gürtels gibt es auch solche, die der Sonne viel näher kommen. Besonders interessant sind dabei die *Apolloasteroiden*, die die Erdbahn kreuzen. Der engste Vorübergang eines Planetoiden an der Erde fand 1937 statt, als Hermes die Erde in nur 680000 km Abstand passierte. Auch weit jenseits des Gürtels scheint es Planetoiden zu geben. So wurde z. B. 1977 der Planetoid 2060 Chiron entdeckt, der die Sonne zwischen den Bahnen von Saturn und Uranus umläuft.
Mit Ausnahme der großen Planetoiden, die Kugelgestalt besitzen, sind die meisten dieser Himmelskörper irregulär begrenzte Brocken. Ihr Aussehen dürfte dem der Marsmonde Phobos und Deimos gleichen. Die *spektralen Eigenschaften* des an ihnen reflektierten Sonnen-

lichts zeigen, daß einige Planetoiden aus ähnlichem Material bestehen wie die Meteorite. Während die großen Planetoiden auf ähnlichem Wege entstanden sein dürften wie die Planeten (↑ 10.2.3.15)., entstanden die zahlreichen kleineren durch die *Zertrümmerung* größerer infolge Zusammenstoßes. Auch die auf die Erde fallenden Meteorite sind wahrscheinlich solche Zertrümmerungsprodukte (Tab. 10.2.3-11).

Tab. 10.2.3-11 Eigenschaften einiger kleiner Planeten; *1* Objekte des Planetoidengürtels, *2* Vertreter der Apollo-Asteroiden, die die Erdbahn kreuzen, *3* Vertreter der Trojaner, die sich in der Jupiterbahn bewegen, *4* erster Vertreter einer neuen Gruppe äußerer Planetoiden

Objekt	Große Bahnhalbachse in AE	Numerische Exzentrizität	Bahnneigung in °	Durchmesser in km	Masse in 10^{21} kg	Bemerkungen
1 Ceres	2,767	0,08	10,6	1003	1,17	*1*
2 Pallas	2,769	0,24	34,8	608	0,26	*1*
3 Juno	2,670	0,26	13,0	247	0,20	*1*
4 Vesta	2,362	0,09	7,1	538	0,24	*1*
433 Eros	1,458	0,24	10,8	23		*2*
1566 Ikarus	1,078	0,83	23	1		*2*
– 1976 UA	0,844	0,45	5,8			*2*
588 Achilles	5,21	0,15	10,3	70		*3*
2060 Chiron	13,66	0,38	7			*4*

10.2.3.13. Kometen. Zu den kleinen Himmelskörpern des Sonnensystems gehören auch jene festen Körper, die den seit urdenklichen Zeiten die Menschen beeindruckenden Kometenerscheinungen zugrunde liegen. Erst in Sonnennähe umgibt sich ein solcher *Kometenkern* durch das Verdampfen seiner flüchtigen Bestandteile mit einer Hülle leuchtenden Gases, der *Koma*, und bildet in den meisten Fällen auch einen *Schweif* aus, der dann für die spektakuläre Kometenerscheinung sorgt (Abb. 10.2.3-12, Tafel 62).

Nach ihren Bahnen definiert man zwei Klassen von Kometen. Die *kurzperiodischen Kometen*, zu denen ungefähr 100 Objekte mit Umlaufzeiten kleiner als 200 Jahre gehören, bewegen sich auf stark elliptischen Bahnen im Abstandsbereich der Planeten um die Sonne. Die *langperiodischen Kometen* – in historischer Zeit wurden etwa 500 derartige Objekte entdeckt – benötigen mehr als 200 Jahre für einen Umlauf um die Sonne, in den meisten Fällen Jahrtausende oder sogar Jahrmillionen. Sie bewegen sich auf parabelnahen Ellipsenbahnen, deren Ebenen beliebige räumliche Orientierung zeigen, und befinden sich die meiste Zeit ihres Bahnumlaufs weit außerhalb des eigentlichen Planetensystems. Wahrscheinlich sind die kurzperiodischen Kometen aus langperiodischen hervorgegangen, deren Bahnen unter der störenden Wirkung der Planeten, vor allem des Jupiters, drastisch verkleinert wurden.

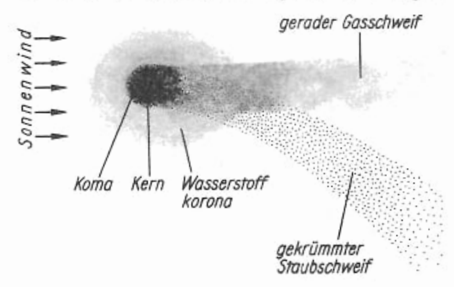


Abb. 10.2.3-12 Erscheinungsbild eines Kometen in schematischer Darstellung

Aus der Häufigkeit ihres Auftretens im inneren Sonnensystem folgt, daß es im Sonnenumkreis von etwa einem Lichtjahr ein großes Reservoir an Kometen geben muß. Zu dieser *Kometenwolke* gehören mindestens einige Millionen, wenn nicht sogar viele Milliarden Objekte. Bei den meisten von ihnen dürfte es sich um Gebilde handeln, die von Anfang an Mitglieder des Sonnensystems waren und nicht etwa aus dem interstellaren Raum eingefangen wurden.

Im *Spektrum der Koma* wurden die Emissionslinien zahlreicher Radikale, z. B. C_2, C_3, CH, CN, NH, NH_2, OH, gefunden, die als Dissoziationsprodukte von Molekülen der Elemente C, N, O und H aufgefaßt werden. Es gelang auch, Moleküle wie H_2O, HCN und CH_3CN nachzuweisen. Die Modellvorstellungen über die chemische Beschaffenheit der *Kometenkerne* gehen davon aus, daß der Kometenkern eine Art Ball aus „schmutzigem Schnee" von Kilometergröße darstellt. Hauptbestandteil sind gefrorene flüchtige Verbindungen wie Wasser (H_2O), Ammoniak (NH_3), Methan (CH_4), Dicyan (C_2N_2), Methylcyanid (CH_3CN), Kohlendioxid (CO_2). In diese Matrix aus Schnee

bzw. Eis sind feste Partikel aus Silicatgestein eingebettet.
Durch die Wechselwirkung des Sonnenwinds (↑ 10.2.3.14.) mit dem abströmenden ionisierten Gas der Koma kommt es zur Ausbildung des geraden von der Sonne weg gerichteten *Gasschweifs*, in dem Molekülionen wie CO^+, CO_2^+ und N_2^+ nachgewiesen wurden. Die gleichfalls aus dem Kern freigesetzten festen Teilchen bilden den gekrümmten *Staubschweif*. Bei vielen Kometen wurden in Sonnennähe Auflösungserscheinungen beobachtet. Es kam dabei zur Aufspaltung in Teilstücke oder zum Erlöschen des Kometen infolge völligen Zerfalls. Die Auflösungsprodukte verteilten sich dann längs der Bahn. Kreuzt die Erde die Bahn eines solchen Meteorstroms, dann treten gehäuft *Sternschnuppenfälle* auf.

10.2.3.14. Interplanetares Medium. Die wichtigste Komponente im Raum zwischen den Planeten ist das ständig von der Sonne abströmende *Plasma des Sonnenwinds*. Dieses etwa 160000 K heiße Gas besteht hauptsächlich aus ionisiertem Wasserstoff und Helium und strömt in Erdbahnnähe mit Geschwindigkeiten zwischen 350 und 800 kms^{-1}. In Erdnähe beträgt seine Dichte im Mittel $5{,}4 \cdot 10^{-6}$ Ionen pro m^3. Die in diesem Plasma eingefrorenen Magnetfelder wechselwirken mit den planetaren Magnetfeldern und erzeugen die charakteristische Struktur der Magnetosphären, z. B. die *Bugstoßwelle*, die *Magnetopause*, den *Magnetschweif*. In großer Entfernung von der Sonne kommt es zur Wechselwirkung mit dem interstellaren Gas und den interstellaren Magnetfeldern. Dadurch wird die Einflußsphäre des solaren Magnetfelds, die *Heliosphäre*, begrenzt. Neben dem Sonnenwindplasma bewegen sich auch Ströme schneller Teilchen durch den interplanetaren Raum, die aus Aktivitätsgebieten der Sonne stammen und auf die planetarische Magnetosphären und Kometenschweife empfindlich reagieren (s. a. 10.2.2.3.).
Die *staubförmige Komponente* des interplanetaren Mediums bildet eine die Sonne einhüllende Wolke, deren Symmetrieebene die Ekliptik ist. Durch die Streuung des Sonnenlichts an diesen Partikeln entsteht die Erscheinung des *Zodiakallichts*, die in Sonnennähe in die *F-Korona* übergeht. In Form von Mikrometeoriten können diese Teilchen mit speziellen Detektoren an Bord von Erdsatelliten und Raumsonden nachgewiesen werden. Neben diesen kleinen Teilchen mit Durchmessern von 10^{-7} bis 10^{-4} m gibt es auch in wesentlich geringerem Umfang größere Brocken im interplanetaren Raum. Alle diese Teilchen und Brocken werden als Auflösungsprodukte von Kometen und als Ergebnisse von Kollisionen kleiner Planeten aufgefaßt und *Meteoroide* bzw. *Mikrometeoroide* genannt. Beim Eindringen in die Erdatmosphäre erzeugt ein Meteoroid eine Meteorerscheinung, z. B. eine *Sternschnuppe* oder eine *Feuerkugel*. Nur hinreichend große und feste Brocken überstehen den Durchflug der Atmosphäre und werden als *Meteorite* am Erdboden gefunden. Auch bestimmte Mikrometeoroide können die Atmosphäre durchqueren, weil sie bereits in großen Höhen so stark abgebremst werden, daß sie nicht mehr verglühen können. Man findet sie auf arktischen Schneeflächen und im Tiefseeschlamm als Mikrometeorite.

10.2.3.15. Entstehung des Planetensystems. Eine geschlossene Theorie der Entstehung des Sonnensystems gibt es bis jetzt noch nicht. In den 300 Jahren wissenschaftlicher Beschäftigung mit den Planeten wurden von zahlreichen bedeutenden Astronomen, Physikern, Mathematikern und Vertretern der Geowissenschaften mehr als 50 Hypothesen unterbreitet, die dieses kosmogonische Problem lösen sollten. Viele davon erwiesen sich als unhaltbar, manche enthielten jedoch sehr fruchtbare Ideen, auf denen die heutige Forschung aufbaut, z. B. die Konzeption des *Sonnennebels*, die Vorstellung der *gemeinsamen Entstehung* von Sonne und Planeten, die Konzeption, daß die Planeten aus *Kondensationsprodukten* des Sonnennebelgases, den *Planetesimals*, agglomerierten.
Planetenkosmogonie ist heute ein interdisziplinäres Forschungsgebiet, in dem die neuen Erkenntnisse der Planetenwissenschaften des Raumflugzeitalters und die Erkenntnisse der Sternkosmogonie tonangebend sind. Die Entstehung des Planetensystems sieht man heute nicht mehr als ein isoliertes Forschungsproblem, sondern man ordnet diese kosmogonischen Prozesse in eine gesetzmäßige Folge von Ereignissen, die mit der Sternentstehung verknüpft sind, ein.
Der früher ad hoc postulierte flache rotierende *Urnebel* ergibt sich als eine zwangsläufige Erscheinung bei der Kontraktion eines instabilen rotierenden Fragments einer interstellaren Wolke. Die physikalischen und chemischen Bedingungen innerhalb dieser dichten Nebelscheibe, die man auf dem Wege über Modellrechnungen erschließt, sind so beschaffen, daß sich im Zentrum eine starke Massenkonzentration bildet, während in den kühleren, vom Zentrum weiter entfernten Gebieten jene Stoffe kondensieren, die das Baumaterial der Himmelskörper bilden. Dabei gibt es allerdings noch zahlreiche Prozesse, die man nur qualitativ versteht. Zu diesen Schwerpunktfragen der heutigen Planetenkosmogonie gehören z. B., wie der große Betrag des Drehimpulses der entstehenden Sonne an die umgebenden Kondensate übertragen wurde, welche Strömungs- und Transportprozesse im einzelnen im Sonnen-

nebel abliefen, wie schnell die Kondensate wuchsen und welche innere Struktur sie besaßen, wie aus den Planetesimals Planeten wurden, welche Stellung die heute noch existierenden Kleinkörper einnehmen, wie sich die Satellitensysteme bildeten usw. Zentrales Anliegen der Bemühungen um die Planetenkosmogonie ist heute die Aufklärung der Herkunft und der Frühgeschichte unseres Heimatplaneten.

10.3. Erscheinungsformen der Materie im Milchstraßensystem

10.3.1. Zustandsgrößen der Sterne

Der physikalische Zustand eines Sterns als Ganzes kann durch Größen charakterisiert werden, die der direkten Beobachtung zugänglich sind, aber nur für sehr wenige Sterne gelingt es, ein vollständiges System dieser Zustandsgrößen zu gewinnen.

10.3.1.1. Sternmasse. Die Masse eines Sterns läßt sich aus der beobachteten Gravitationswirkung bestimmen, z. B. erhält man für die Massensumme der Komponenten eines *visuellen Doppelsterns* nach dem 3. Keplerschen Gesetz $M_A + M_B = a^3 : U^2$ in Vielfachen der Sonnenmasse, wenn a die große Halbachse der Bahnellipse des Begleiters B um den Hauptstern A in astronomischen Einheiten ist und U die Umlaufzeit in Jahren bedeutet (Abb. 10.3.1-1 a). Die große Halbachse a läßt sich aber nur im linearen Maß angeben, wenn außer dem Winkelabstand der beiden Komponenten auch die Entfernung des Doppelsterns bekannt ist. Bei unbekannter Entfernung des Doppelsterns läßt sich a aus der Länge der Bahn des Begleiters B berechnen, wenn neben der Umlaufzeit U seine Bahngeschwindigkeit aus spektroskopischen Beobachtungen bestimmt werden konnte. Die Masse jeder einzelnen Komponente erhält man nur dann, wenn die Bewegung beider

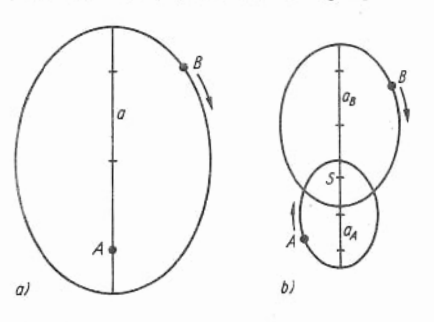


Abb. 10.3.1-1 Visueller Doppelstern; a) die relative Bahn der Komponente B in bezug auf A, b) die Komponenten A und B bewegen sich um den gemeinsamen Schwerpunkt S

Sterne um den gemeinsamen Schwerpunkt S bekannt ist (Abb. 10.3.1-1 b).

Genaue Massenangaben liegen bisher nur für rund 40 visuelle Doppelsterne und 20 Bedeckungsveränderliche vor, bei denen sich die Bahngeschwindigkeiten spektroskopisch ergeben haben. Der Stern mit der geringsten bis jetzt gemessenen Masse ist Luyten 726-8 B mit nur 0,04 Sonnenmassen. Die obere Grenze liegt nach theoretischen Überlegungen bei 60 Sonnenmassen.

Im Prinzip kann die Masse eines Einzelsterns bei bekanntem Radius auch aus der *Gravitationsrotverschiebung* seiner Spektrallinien bestimmt werden, was aber nur in den seltensten Fällen möglich ist.

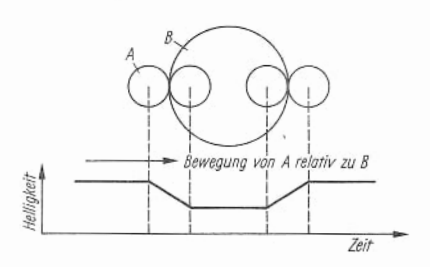


Abb. 10.3.1-2 Bedeckung der Komponente A durch B und schematische Lichtkurve

10.3.1.2. Sterndurchmesser. Der Winkeldurchmesser ist nur bei der Sonne direkt mit einem Mikrometer meßbar; alle anderen Sterne sind zu weit entfernt dafür. Bei einigen nahen Riesensternen kann der Winkeldurchmesser auch interferometrisch bestimmt werden.

Die fotometrischen und geometrischen Verhältnisse bei *Bedeckungsveränderlichen*, bei denen sich zwei Sterne so umeinander bewegen, daß sie sich abwechselnd gegenseitig verdecken, gestatten eine relativ genaue Bestimmung ihrer Durchmesser, wenn aus spektroskopischen Beobachtungen die Bahngeschwindigkeit und durch die Beobachtung des Lichtwechsels die Umlaufzeit bekannt ist (Abb. 10.3.1-2).

Die *strahlungstheoretische Methode* benutzt die Beziehung

(1) $$D^2 = L/(\pi\sigma T_{\text{eff}}^4)$$

zwischen dem Durchmesser D, der Leuchtkraft L und der effektiven Temperatur T_{eff} des Sterns, in der σ die Stefan-Boltzmann-Konstante bedeutet.

Den kleinsten Durchmesser, nur wenige Kilometer, besitzen die *Neutronensterne*, ihnen folgen die *Weißen Zwerge*, z. B. Wolf 219 mit 0,36 Erddurchmessern. *Überriesen*, wie VV Cephei mit 2400 Sonnendurchmessern, erreichen das Ausmaß von Planetenbahnen.

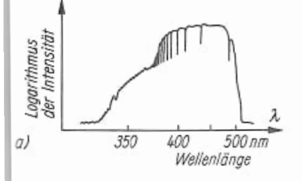
10.3.1.3. Leuchtkraft. Die Leuchtkraft L eines Sterns gibt die gesamte von ihm je Sekunde ausgestrahlte Energie an. Sie läßt sich nach $L = 2{,}72 \cdot 10^{35-0{,}4\,M_{bol}}$ aus der in Größenklassen gemessenen *absoluten bolometrischen Helligkeit* M_{bol} berechnen; ein Stern mit $M_{bol} = 0^m$ strahlt danach pro Sekunde die Energie $2{,}72 \cdot 10^{35}$ erg $= 2{,}72 \cdot 10^{28}$ J aus. Der Wert M_{bol} wiederum kann aus der *scheinbaren Helligkeit* m und der in Parsec gemessenen Entfernung r nach der Beziehung $M = m + 5 - 5 \cdot \log r$ ermittelt werden, die auf der Annahme beruht, daß das Entfernungsquadratgesetz der Lichtausbreitung gilt, also keine interstellare Extinktion vorhanden ist. Ist die Entfernung r vom Beobachter unbekannt, so läßt sich die Leuchtkraft L nach Leuchtkraftkriterien aus dem *Sternspektrum* (Tafel 60) ermitteln. Die Leuchtkraft zweier Sterne mit gleicher effektiver Temperatur ist um so größer, je größer ihr Durchmesser ist; von diesem hängen aber der Druck und damit der Ionisationsgrad in der Sternatmosphäre ab und deshalb auch die Stärke bestimmter Spektrallinien. Man kann daher aus der Stärke dieser Spektrallinien auf die Leuchtkraft eines Sterns schließen. Auch die Breite bestimmter Spektrallinien gibt Aufschluß über seine Leuchtkraft: *Überriesen* haben wegen der geringen Schwerebeschleunigung an ihrer Oberfläche und der geringen Druckverbreiterung schärfere Linien als die *Zwergsterne* gleicher effektiver Temperatur.

Bei den Überriesen kann die Leuchtkraft um einen Faktor 10^5 größer als die der Sonne sein. Die Leuchtkraft von Wolf 1055 beträgt hingegen nur das $6{,}6 \cdot 10^{-5}$fache der Sonnenleuchtkraft.

10.3.1.4. Effektive Temperatur. Die Temperatur eines Sterns kann nur für seine Atmosphäre aus den Beobachtungen direkt abgeleitet werden, weil nur aus dieser obersten Schicht beobachtbare Strahlung stammt. Da die Sternatmosphären nicht im thermodynamischen Gleichgewicht sind, strahlen sie auch nicht wie schwarze Körper (Abb. 10.3.1-3); unterschiedliche Meßverfahren führen daher im allgemeinen zu unterschiedlichen Temperaturangaben. Unter der effektiven Temperatur T_{eff} eines Sterns versteht man die Temperatur, die ein schwarzer Körper haben müßte, der je Flächen- und Zeiteinheit die gleiche Energiemenge ausstrahlt wie der Stern an seiner Oberfläche. T_{eff} kann über die Beziehung 10.3.(1) berechnet werden. Während aber L als die Ausstrahlung über das gesamte Spektrum festgelegt ist, wird auf der Erde nur der von ihrer Atmosphäre nicht absorbierte Teil gemessen; man erhält keine wahren effektiven Temperaturen, sondern Werte, die kleiner als die wahren sind. Um diese Schwierigkeiten zu umgehen, definiert man die *Strahlungstemperatur* als die Temperatur, die ein schwarzer Strahler haben müßte, der im gleichen Spektralbereich je Flächen- und Zeiteinheit die gleiche Energiemenge ausstrahlt wie der Stern, und wählt diesen Spektralbereich so aus, daß er ganz im Durchlässigkeitsbereich der Erdatmosphäre liegt. Vergleicht man vom Stern und von einem schwarzen Körper nicht die ausgestrahlten Energiemengen, sondern nur die beiden Intensitätsverteilungen, d. h. die Form der Energiekurven, so erhält man eine *Farbtemperatur*. Je nach dem benutzten Farbbereich ergeben sich unterschiedliche Temperaturwerte.

Die Farbtemperatur ist am leichtesten durch Beobachtungen zu bestimmen. Der physikalische Zustand einer Sternatmosphäre wird aber durch die effektive Temperatur noch am besten charakterisiert. Mit Hilfe von Sternen, bei denen man sowohl die effektive Temperatur als auch Farbtemperaturen ableiten kann, lassen sich empirische Beziehungen herleiten, die eine Umrechnung der Temperaturwerte möglich machen. Die effektiven Temperaturen der Sternatmosphären liegen im allgemeinen zwischen rund 2500 K und 50000 K.

10.3.1.5. Spektraltyp. Die Spektralklasse, der Spektraltyp, charakterisiert die Art eines Sternspektrums, in dem einem *kontinuierlichen Spektrum* mehr oder weniger viele Absorptions- und Emissionslinien überlagert sind. Im kontinuierlichen Spektrum hängt die Energieverteilung von der effektiven Temperatur ab; das Maximum der Energiestrahlung verschiebt sich mit wachsender Temperatur zu kürzeren Wellenlängen. Die Stärke der *Spektrallinien* wird einerseits durch die Häufigkeit der Elemente in der Sternatmosphäre bedingt, andererseits durch den Ionisationsgrad und den Anregungszustand der Atome und Ionen dieses Elements, die mit steigender Temperatur zu-



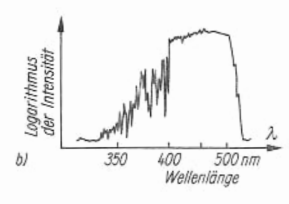


Abb. 10.3.1-3 Intensitätsverteilung a) im Spektrum eines A0-Sterns und b) eines K5-Sterns

nehmen; mit steigendem Druck fällt außerdem der Ionisationsgrad.

In den Sternspektren spiegeln sich somit physikalische Bedingungen wie Temperatur oder Schwerebeschleunigung und unterschiedliche chemische Zusammensetzung der Sternatmosphäre wider.

Da aber die chemische Zusammensetzung im allgemeinen wenig von Stern zu Stern variiert, werden die Sterne durch eine Spektralklassifikation hauptsächlich nach den physikalischen Bedingungen geordnet. Wegen der zwei unabhängigen Parameter, der effektiven Temperatur und der Schwerebeschleunigung, führt man für die Sternspektren im allgemeinen eine *zweiparametrige Spektralklassifikation* durch. Neben der Spektralklasse gibt man noch die Leuchtkraftklasse eines Sterns an. Sterne gleicher *Spektralklasse* haben etwa gleiche *effektive Temperatur*; die *Leuchtkraftklasse* stellt dann ein Maß für die in der Sternatmosphäre herrschende *Schwerebeschleunigung* dar. Betrachtet man Sterne gleicher Leuchtkraftklasse, so ist der Spektraltyp ein Maß für die effektive Temperatur.

Die Hauptfolge der *Spektralklassen* wird durch die Buchstaben W, O, B, A, F, G, K, M gekennzeichnet, die gleichzeitig eine Folge sinkender effektiver Temperaturen ist; zu Nebenfolgen gehören die Spektralklassen R und N bzw. S. Zur feineren Unterscheidung unterteilt man die Spektralklassen noch dezimal mit Ausnahme der Spektralklassen W, R, N und S durch angefügte Ziffern 0 bis 9; im Sinne fallender effektiver Temperatur folgt z. B. auf A9 die Spektralklasse F0. Es hat sich eingebürgert, die Spektralklassen W bis A als frühe, die von F bis G als mittlere und die übrigen als späte Spektralklassen zu bezeichnen, ohne daß damit eine kosmogonische Vorstellung verknüpft wird, wie es früher geschah.

Aus einer Folge typischer Sternspektren verschiedener Spektralklassen (Tafel 60) ist zu erkennen, wie z. B. die Stärke der Linien der Balmerserie H_β, H_γ, H_δ, die vom neutralen Wasserstoff absorbiert werden, von B0 über B5 bis zu einem Maximum bei A0 wächst und dann mit weiter sinkender effektiver Temperatur schwächer wird, während diese Linien bei den späten Spektralklassen unter der Vielzahl der Metallinien kaum noch in Erscheinung treten. Dies ist auf keinen Fall die Folge unterschiedlicher Häufigkeit des Wasserstoffs, sondern allein durch den unterschiedlichen Anregungszustand der Wasserstoffatome bei den verschiedenen effektiven Temperaturen bedingt. Als allgemeine Regel gilt, daß um so mehr ionisierte Atome am Linienspektrum beteiligt sind, je höher die effektive Temperatur ist, und um so mehr Moleküle, je niedriger sie ist.

Einige Besonderheiten in einzelnen Spektren werden durch kleine Buchstaben gekennzeichnet, die man den Großbuchstaben und Ziffern hinzufügt; diffus verwaschene (neblige) Spektrallinien werden durch ein nachgestelltes n, scharfe durch s und Emissionslinien durch e markiert. Zeigt das Spektrum die Merkmale eines Riesensterns (Gigant) oder eines Zwergsterns [engl. *dwarf*], so wird ein g oder ein d vorgesetzt. Die Sonne ist danach ein dG1-Stern.

Die Kennzeichnungen mit d und g stellen Vorläufer der jetzt üblichen Einteilung der Sternspektren in *Leuchtkraftklassen* I bis VI dar; I bedeutet *Überriesen*, unterteilt in Ia *helle* und Ib *schwächere* Überriesen, II *helle Riesen*, III *normale Riesen*, IV *Unterriesen*, V *Hauptreihensterne* oder *Zwerge*, VI *Unterzwerge*. Nach dieser Bezeichnung ist die Sonne ein G1V-Stern.

Bei Sternen geringer scheinbarer Helligkeit ist eine Spektralklassifikation nur sehr schwer möglich. Als Ersatz für die Spektralklasse gibt man dann den *Farbenindex* des Sterns an, das ist die Differenz zwischen zwei in verschiedenen Bereichen des Spektrums (Farben) gemessenen scheinbaren Helligkeiten des Sterns, z. B. zwischen Blauhelligkeit und Gelbhelligkeit. Als Farbenbereich benutzt man meist das international festgelegte *UBV-System*, in dem eine Ultraviolettthelligkeit U, eine Blauhelligkeit B und eine Helligkeit V im visuellen Bereich gemessen und die Farbenindizes $U-B$ und $B-V$ gebildet werden.

10.3.1.6 Magnetfeld. Das Magnetfeld eines Sterns macht sich durch eine Aufspaltung, zumindest aber durch eine Verbreiterung seiner Spektrallinien nach dem Zeemaneffekt bemerkbar (↑ 4.3.3.3.3.). Da der Effekt sehr klein ist, beschränken sich die Beobachtungsmöglichkeiten auf Sterne mit wenigen, aber scharfen Spektrallinien, bei denen eine Linienverbreiterung weder durch schnelle Rotation noch durch große Schwerebeschleunigung oder starke Turbulenz in der Sternatmosphäre verursacht wird. Die gemessenen Flußdichten betragen bis zu 3,4 T und sind vielfach nicht konstant, sondern veränderlich; es wurden sogar Umpolungen beobachtet; z. B. schwankt das Magnetfeld des Sterns 53 Camelopardalis innerhalb von 8 Tagen zwischen $+0{,}375$ und $-0{,}539$ T.

Da großräumige stellare Magnetfelder dieser Stärke in der elektrisch gut leitenden Sternmaterie „eingefroren" sind und nicht innerhalb derart kurzer Zeiten zerfallen bzw. sich umpolen können, wird zur Deutung der Beobachtungen ein schiefes Rotatormodell herangezogen: magnetische und Rotationsachse fallen nicht zusammen, so daß wir abwechselnd magnetische Gebiete unterschiedlicher Stärke und Polarität beobachten. Ungeklärt ist bis heute das Auf-

treten von Linien Seltener Erden, z. B. Europium, und anderer sonst nicht beobachteter Elemente in den Atmosphären einiger magnetischer Veränderlicher. Die Intensitäten dieser Linien sind häufig ebenfalls veränderlich.

10.3.1.7. Rotation. Die Rotation eines Sterns macht sich durch eine Linienverbreiterung im Spektrum bemerkbar (Abb. 10.3.1-4). Das Gesamtspektrum ist die Überlagerung der

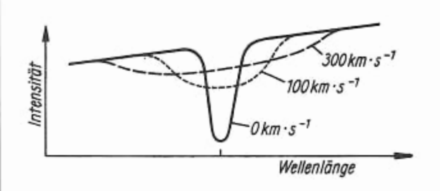


Abb. 10.3.1-4 Schematische Darstellung der Verbreiterung von Spektrallinien durch Sternrotation

Spektren einzelner Punkte auf der Sternoberfläche, die in bezug auf den Beobachter unterschiedliche Radialgeschwindigkeit und damit nach dem *Dopplereffekt* verschiedene Verschiebung der Spektrallinien haben. Das Ergebnis ist eine symmetrische Verbreiterung der Spektrallinien, aus der sich auf die Rotationsgeschwindigkeit schließen läßt. Die erhaltenen Werte stellen Minimalwerte dar; nur wenn die Rotationsachse senkrecht zur Sichtlinie Beobachter – Stern liegt, erhält man die wahre Rotationsgeschwindigkeit.

Bei Bedeckungsveränderlichen läßt sich eine vorhandene Rotation dann gut feststellen, wenn eine kleine Komponente *A* mit hoher Flächenhelligkeit durch eine große Komponente *B* geringerer Flächenhelligkeit verdeckt wird (Abb. 10.3.1-5). Unmittelbar vor und nach der Bedeckung der Komponente *A* bleibt ein schmaler Randbezirk frei, der sich bei Rotation vor der Bedeckung vom Beobachter weg, nach ihr aber auf ihn zu bewegt, wenn Rotation und Bahnbewegung im gleichen Sinn erfolgen.

Im allgemeinen ist die Rotationsgeschwindigkeit für Sterne früher Spektralklassen höher als für die mittleren Spektraltyps und relativ hoch für Oe- und Be-Sterne, in deren Spektren Emissionslinien auftreten. Wahrscheinlich verursachen bei Oe- und Be-Sternen Instabilitäten am Äquator das Abströmen leuchtender Gasmassen, die den Stern umgeben und dadurch das Auftreten der Emissionslinien erklären

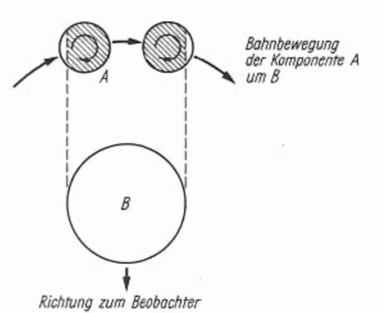


Abb. 10.3.1-5 Geometrische Verhältnisse bei einem Bedeckungsveränderlichen unmittelbar vor und nach der Bedeckung einer Komponente durch die andere

Tab. 10.3.1-6 Mittlere Zustandsgrößen für verschiedene Spektral- und Leuchtkraftklassen *SLK* in Einheiten der Sonnenwerte, *L* Leuchtkraft; *M* Masse, *D* Durchmesser, T_{eff} effektive Temperatur, ϱ mittlere Dichte, *g* Schwerebeschleunigung an der Oberfläche; für die Sonne als GIV-Stern gilt $L = 3{,}86 \cdot 10^{26}$ W, $M = 1{,}99 \cdot 10^{33}$ g, $D = 1{,}4 \cdot 10^{11}$ cm, $T_{eff} = 5785$ K, $\varrho = 1{,}4$ g cm^{-3}, $g = 2{,}7 \cdot 10^{4}$ cm s^{-2}

SLK		*L*	*M*	*D*	T_{eff}	ϱ	*g*
Hauptreihen-sterne	BO V	8000	15	9,0	3,8	0,02	0,2
	AO V	60	3	2,2	1,8	0,3	0,6
	FO V	6	1,5	1,5	1,3	0,5	0,7
	GO V	1	1,0	1,0	1,0	1,0	1,0
	KO V	0,4	0,8	0,8	0,8	1,6	1,2
	MO V	0,06	0,5	0,6	0,6	2	1,4
Riesen	FO III	15	2,5	4	1,0	0,04	0,2
	GO III	40	3,1	10	0,9	0,003	0,03
	KO III	80	3,5	25	0,7	0,0002	0,006
	MO III	400	3,8	75	0,5	0,000009	0,0007
Überriesen	BO I	200000	50	20	4,7	0,006	0,1
	AO I	20000	15	40	1,9	0,0002	0,009
	FO I	6000	13	65	1,1	0,00005	0,003
Weiße Zwerge	AO	0,0005	0,6	0,02	1,2	70000	1500

(Tab. 10.3.1-6). Die größten bis jetzt für die Oberfläche von Sternen gefundenen Rotationsgeschwindigkeiten betragen 200 bis 400 kms^{-1}. Bei bekanntem Durchmesser der Sterne ergeben sich *Rotationsperioden* bis herab zu 0,6 Tagen.

10.3.1.8. Chemische Zusammensetzung der Sternatmosphären. Über die Spektralanalyse des von einem Stern ausgesandten Lichts läßt sich die chemische Zusammensetzung der Schichten des Sterns bestimmen, von denen das Licht emittiert wird. Man kann daher die Häufigkeit der Elemente nur in den Sternatmosphären direkt untersuchen (↑ 10.3.2.).

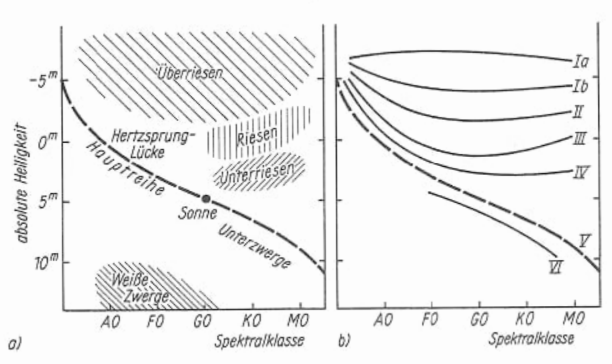


Abb. 10.3.1-7 Schematisches Hertzsprung-Russell-Diagramm mit a) der Lage einiger wichtiger Sterngruppen und b) der Anordnung der Leuchtkraftklassen

Das weitaus häufigste Element ist der Wasserstoff. Ihm folgen Helium und die schwereren Elemente, deren Masseanteil aber maximal einige Prozent nicht übersteigt. Je größer die Massenzahl ist, um so seltener ist im allgemeinen das Element.

Detailliertere Untersuchungen zeigen, daß sich die Sterne in dem Häufigkeitsverhältnis Wasserstoff zu Metallen zum Teil merklich unterscheiden. Als Metalle werden hier alle schweren Elemente, vom Kohlenstoff an, bezeichnet. Die Sterne mit sonnenähnlicher Atmosphärenchemie werden gewöhnlich unter dem Begriff der *Population I* zusammengefaßt, alle metallarmen Sterne gehören der *Population II* an.

Anomale Elementenhäufigkeit findet sich bei einigen recht seltenen Sterntypen; z. B. kennt man Sterne, die anomal viel Helium, dagegen sehr wenig Wasserstoff haben. Andere Sterne zeichnen sich durch Überhäufigkeiten von Kohlenstoff, Stickstoff oder sogar bestimmten Schwermetallen, wie Zirconium, Yttrium, Niobium, Molybdän, Barium und den Seltenen Erden, aus.

Die *Isotopenhäufigkeiten* der meisten Elemente weichen kaum von den irdischen Häufigkeitsverhältnissen ab, jedoch gibt es auch bemerkenswerte Ausnahmen, z. B. die relative Überhäufigkeit des Kohlenstoffisotops C-13 in den meisten Kohlenstoffsternen. Für andere beobachtete Überhäufigkeiten, wie im Falle von He-3 und H-3 scheinen Kernreaktionen in den Sternatmosphären verantwortlich zu sein. Nach starken Sonneneruptionen tritt beispielsweise H-3 verstärkt im Sonnenwind (↑ 10.1.4.1.) auf. Die in einigen magnetischen Sternen beobachtete Überhäufigkeit von He-3 kann mit den starken lokalen Magnetfeldern dieser Sterne zusammenhängen, die als Energiequelle für derartige Eruptionen in Frage kommen.

10.3.1.9. Zustandsdiagramme. Zwischen einzelnen Zustandsgrößen (↑ Tab. 10.3.1-6) bestehen formelmäßig ausdrückbare Beziehungen, z. B. wird die Leuchtkraft eines Sterns durch seine effektive Oberflächentemperatur und die Größe seiner Oberfläche und damit vom Sterndurchmesser bestimmt. Die Beziehungen zwischen anderen Zustandsgrößen, z. B. zwischen Spektralklasse, effektiver Temperatur und Schwerebeschleunigung lassen sich nicht formelmäßig ausdrücken; sie werden durch Zustandsdiagramme verdeutlicht.

10.3.1.9.1. *Hertzsprung-Russell-Diagramm.* Im Hertzsprung-Russell-Diagramm trägt man über der Spektralklasse die absolute Helligkeit oder die Leuchtkraft der Sterne auf. Dabei ist nicht jede Kombination dieser zwei Zustandsgrößen in der Natur verwirklicht; das Diagramm ist nicht gleichmäßig mit Sternen besetzt. Diese ordnen sich vielmehr in bestimmten Gebieten, in *Ästen*, an, und andere Gebiete, die *Lücken*, bleiben frei. Die Mehrzahl der Sterne liegt auf dem verhältnismäßig scharf begrenzten *Zwergenast*, auf dem sich die *Hauptreihensterne* der Leuchtkraftklasse V quer durch das Diagramm ziehen, von den O-Sternen mit der absoluten Helligkeit $M = -6^m$ bis zu den M-Sternen mit $M = +16^m$. Ein zweiter, weniger scharf ausgebildeter Ast wird vorwiegend von Sternen der Spektralklassen G^m bis M gebildet; wegen ihrer im Mittel großen absoluten Helligkeit $M = 0^m$ müssen diese Sterne bei nahezu gleicher Spektralklasse und damit gleicher effektiver Temperatur wie die Hauptreihensterne einen größeren Durchmesser als diese haben (Abb. 10.3.1-7).

Sie gehören zur Leuchtkraftklasse III, den *normalen Riesensternen*. Zwischen diesem *Riesenast* und der Hauptreihe liegt das Gebiet der *Unterriesen* der Leuchtkraftklasse IV, das aber nur relativ schwach besetzt ist. Der Riesenast wird von der Hauptreihe durch die *Hertzsprunglücke* getrennt, in der auffallend wenige Sterne liegen. Die Sterne der Leuchtkraftklasse VI, die *Unterzwerge*, befinden sich etwa 1^m bis 3^m unterhalb der Hauptreihe im Bereich der mittleren und späten Spektralklassen. Die *Überriesen* und die *hellen Riesen* der Leuchtkraftklassen I und II nehmen das Gebiet über dem Riesenast und der Hauptreihe ein, sie sind die absolut hellsten Sterne. Ein isoliertes Gebiet etwa 8^m bis 12^m unterhalb der Hauptreihe im Bereich der Spektralklasse B bis G ist von den *Weißen Zwergen* besetzt. Neben diesen Hauptgruppen nehmen noch andere Sterngruppen, deren Mitglieder gleiche oder nahezu gleiche innere Struktur haben, charakteristische Plätze im Hertzsprung-Russell-Diagramm ein, z. B. die Veränderlichen, die Novae.

Für eine Gruppe von Sternen mit gleicher Entfernung vom Beobachter ist bis auf eine konstante Differenz die scheinbare gleich der absoluten Helligkeit. Im Hertzsprung-Russell-Diagramm läßt sich für solche Sterne deshalb die absolute durch die scheinbare Helligkeit ersetzen. Benutzt man außerdem statt der Spektralklasse den Farbenindex, so entsteht ein *Farben-Helligkeits-Diagramm*, in dem die relative Lage der Äste richtig dargestellt ist (Abb. 10.3.1-8).

Zwischen dem Hertzsprung-Russell-Diagramm der Population I und dem der Population II bestehen charakteristische Unterschiede (Abb. 10.3.1-9). Sie führten zur Entdeckung der *Populationen*. Während im Hertzsprung-Russell-Diagramm der Population I die *Hauptreihe* bis zu den B- und O-Sternen besetzt ist, fehlen in dem der Population II die Hauptreihensterne von der Spektralklasse F0 an in Richtung auf die frühen Spektraltypen völlig. Außerdem sind

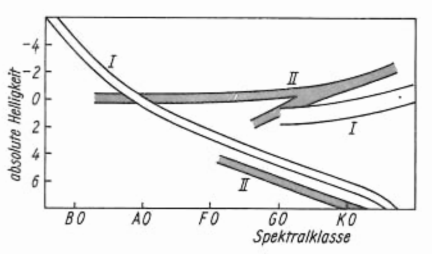


Abb. 10.3.1-9 Schematisches Hertzsprung-Russell-Diagramm für die Populationen I und II

die *Riesenäste* beider Populationen etwas gegeneinander verschoben. Der Riesenast der Population II gabelt sich bei der Spektralklasse G0 in einen horizontalen Ast und in einen, der in Richtung auf die Hauptreihe nach unten läuft. Der Ast der *Unterzwerge* der Leuchtkraftklasse VI ist typisch für die Population II. Charakteristische Beispiele für Hertzsprung-Russell-Diagramme der beiden Populationen liefern die offenen Sternhaufen mit Sternen der Population I und die Kugelsternhaufen mit Sternen der Population II.

Eine theoretische Deutung der Hertzsprung-Rusell-Diagramme hinsichtlich Lage und Besetzungsdichte der verschiedenen Äste gelang in neuerer Zeit aufgrund der Untersuchungen zum inneren Aufbau und zur Entwicklung der Sterne (↑ 10.3.3.).

10.3.1.9.2. *Masse-Leuchtkraft-Beziehung.* Im

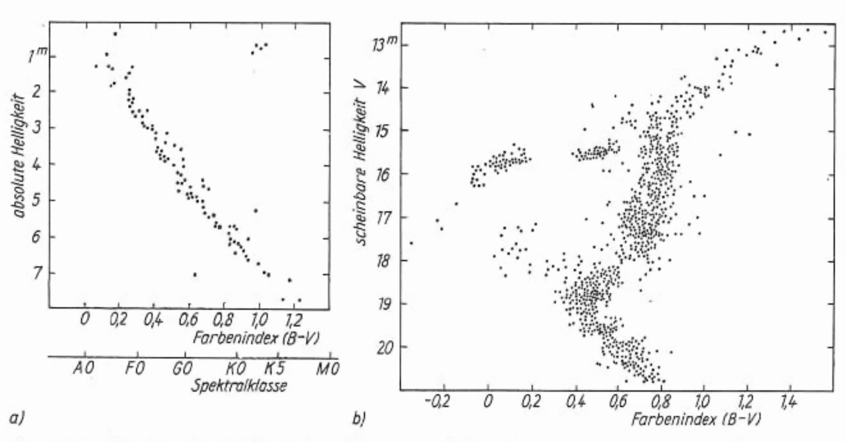


Abb. 10.3.1-8 Farben-Helligkeits-Diagramm a) der Hyaden nach Arp, scheinbare Helligkeiten in absolute umgerechnet; b) des Kugelsternhaufens M 3 nach Johnson und Sandage

Hauptreihenstadium hängt die Leuchtkraft eines Sterns empfindlich von seiner Masse ab. Für Sterne der Population I kann die beobachtete Masse-Leuchtkraft-Beziehung in einem weiten Bereich durch ein Potenzgesetz angenähert werden (Abb. 10.3.1-10). Danach ist die Leuchtkraft ungefähr der 3,5ten Potenz der Masse proportional.

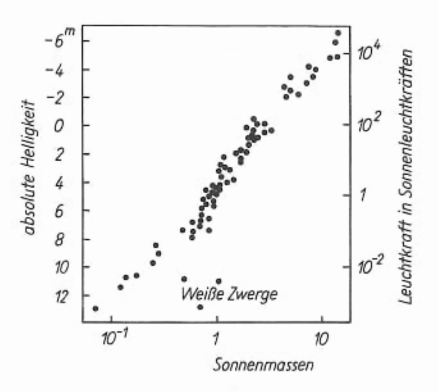


Abb. 10.3.1-10 Beobachtete Masse-Leuchtkraft-Beziehung; die deutlich aus der Folge der Hauptreihensterne herausfallenden drei Sterne sind Weiße Zwerge

10.3.1.10. Sterne mit veränderlichen Zustandsgrößen. Im allgemeinen ist die absolute bolometrische Helligkeit eines Sterns konstant; sie ändert sich nur in großen Zeiträumen infolge der Sternentwicklung. Bei einer großen Anzahl von Sternen, den *Veränderlichen*, aber schwankt die Helligkeit mehr oder minder regelmäßig um einen Mittelwert.

Von den eigentlichen oder *physischen Veränderlichen*, deren Lichtwechsel mit der Änderung physikalischer Zustandsgrößen einhergeht, sind die *optischen Veränderlichen* zu unterscheiden. Bei ihnen handelt es sich um optisch nicht auflösbare Doppelsternsysteme, deren Gesamthelligkeit dadurch variiert, daß periodisch eine Komponente durch die andere verdeckt wird, oder sich die Größe der sichtbaren leuchtenden Sternoberflächen ändert, das ist beispielsweise bei engen Doppelsternen der Fall, deren Komponenten infolge der gegenseitigen Gezeitenwirkung stark verformt sind.

Die Veränderlichen werden gewöhnlich nach ihren *Lichtkurven* klassifiziert. In einer Lichtkurve ist die scheinbare Helligkeit als eine Funktion der Zeit aufgetragen. Aus ihr ergibt sich die Amplitude und gegebenenfalls auch die Periode des Lichtwechsels. Eventuell werden noch weitere Merkmale zur Klassifikation herangezogen, wie absolute Helligkeit, Spektralklasse oder Populationszugehörigkeit. Da aber eine befriedigende Theorie für die Ursachen der Veränderlichkeit fehlt, ist eine physikalisch begründete Klassifikation nicht immer zu finden; es lassen sich bisher nur Teilerscheinungen bei einigen Gruppen von Veränderlichen verstehen.

Rund 95% der physischen Veränderlichen liegen im Hertzsprung-Russell-Diagramm im Gebiet der Riesen und Überriesen, und von ihnen gehört die überwiegende Mehrzahl zu den *Pulsationsveränderlichen*, deren Lichtwechsel durch mehr oder minder regelmäßige Veränderungen des Sterndurchmessers, durch Pulsationen, hervorgerufen wird; gleichzeitig ändert sich dabei die effektive Temperatur. Der Stern führt als Ganzes Schwingungen um eine konstante Mittellage aus. Dies ergibt sich aus den periodischen Verschiebungen der Linien im Spektrum nach Rot bzw. Blau im Rhythmus des Lichtwechsels.

Als Beispiel für die Gruppe der Pulsationsveränderlichen seien die δ-Cephei-Sterne benannt. Die Periode ihres Lichtwechsels liegt zwischen 1 und etwa 50 Tagen; bei einzelnen Sternen kommen kleinere Periodenänderungen vor. Die Helligkeitsschwankungen betragen bis zu $2{,}6^m$ und hängen etwas von der Periodenlänge ab: δ-Cephei-Sterne mit längerer Periode haben auch eine größere Amplitude des Lichtwechsels. Durch die Pulsation ändert sich der Sternradius erheblich, im Durchschnitt um 10%; bei dem Stern δ Cephei selbst, dem Prototyp der Gruppe, ist die Differenz zwischen

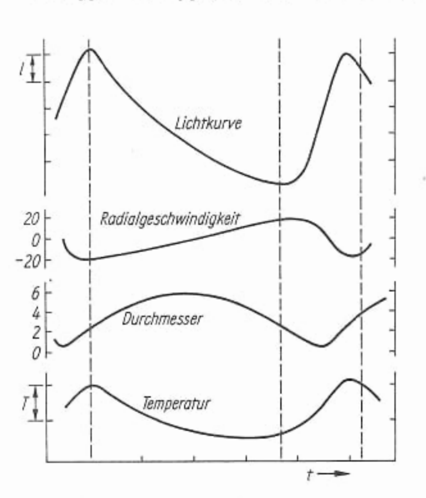


Abb. 10.3.1-11 Verlauf der Radialgeschwindigkeit in km s^{-1}, des Durchmessers in 10^6 km und der Temperatur im Vergleich zur Lichtkurve bei δ-Cephei; Schwankungsbreite der Temperatur $T = 1\,000$ K, der Helligkeit $l = 0{,}2$ m

dem maximalem und den minimalen Radius 2,7 · 10^6 km. Die Kurve der Radialgeschwindigkeiten bei Expansion und Kontraktion ist fast ein Spiegelbild der Lichtkurve; die größte negative Radialgeschwindigkeit, d. h. die größte Expansionsgeschwindigkeit, tritt im Helligkeitsmaximum auf, die größte Kontraktionsgeschwindigkeit im Helligkeitsminimum (Abb. 10.3.1-11). Berechnet man zu jeder Radialgeschwindigkeit den Sternradius, so zeigt sich, daß er für maximale Helligkeit die gleiche Größe hat wie für minimale, daß er während des Abstiegs der Lichtkurve seinen größten und bei zunehmender Helligkeit seinen kleinsten Wert hat. Die Helligkeitsunterschiede zwischen Maximum und Minimum sind daher nicht durch unterschiedliche Größe des Sterns, sondern nur durch die Schwankung der effektiven Temperatur verursacht. Die Temperaturunterschiede betragen rund 1000 K; mit ihnen geht eine Änderung des Spektraltyps einher.

Die *Perioden-Helligkeits-Beziehung* stellt einen physikalischen Zusammenhang zwischen der Periodenlänge und der absoluten Helligkeit her und wird benutzt, die Entfernung der δ-Cephei-Sterne und die von extragalaktischen Sternsystemen zu bestimmen, in denen sie sich befinden (↑ Abb. 10.3.3-8). Wegen ihres regelmäßigen Lichtwechsels und ihrer großen absoluten Helligkeit sind die δ-Cephei-Sterne auch in großer Entfernung zu entdecken; sie gehören zur Population I.

Novae und *Supernovae* werden den *Eruptionsveränderlichen* zugeordnet, die einen plötzlichen Helligkeitsausbruch, der von Materieauswürfen begleitet ist, zeigen. Bei einigen Veränderlichen wiederholen sich derartige Eruptionen mehr oder weniger regelmäßig in jahrelangem Abstand.

Die Novae zeigen innerhalb weniger Stunden einen Helligkeitsanstieg von 7 bis 16 Größenklassen, ihre Ausstrahlung erreicht demzufolge das 10^3- bis 10^6fache des Werts vor dem Ausbruch. Der Helligkeitsrückgang, gelegentlich unterbrochen durch mehr oder minder rasche Helligkeitsschwankungen, erfolgt langsam und dauert Jahre. Die Helligkeit der *Postnova* unterscheidet sich von der der *Praenova* kaum (Abb. 10.3.1-12).

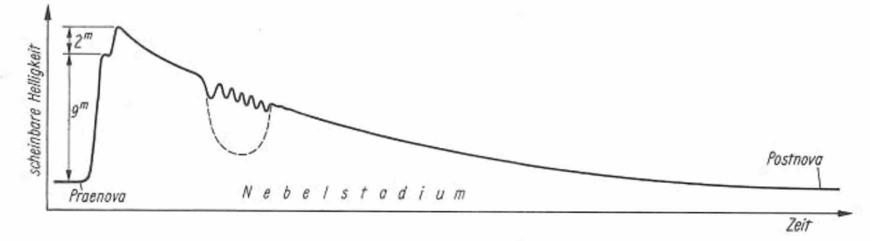


Abb. 10.3.1-12 Schematische Lichtkurve einer Nova

Der Helligkeitsausbruch einer Nova wird dadurch verursacht, daß der Stern seine äußere leuchtende Hülle mit großer Geschwindigkeit abwirft. Aus der beobachteten Dopplerverschiebung der Spektrallinien kann die Expansionsgeschwindigkeit berechnet werden; während des Helligkeitsmaximums werden Werte zwischen rund 600 und 2500 kms^{-1} erreicht. Die Expansionsgeschwindigkeit kann danach sogar noch weiter anwachsen. Der rasche Leuchtkraftanstieg ist einfach eine Folge der zunehmenden Oberfläche der leuchtenden Gashülle. Mit fortschreitender Ausdehnung nimmt die Dichte der ausgeschleuderten Materie immer mehr ab, die beobachteten Absorptionslinien werden schwächer. Im folgenden *Nebelstadium* ist der Stern von außerordentlich dünn verteilter Materie umgeben, in der *verbotene Linien* ausgestrahlt werden können. Im Postnovastadium ist die Dichte der ausgeschleuderten Materie so gering geworden, daß sie nicht mehr nachweisbar ist. Abschätzungen ergeben, daß der Stern bei einem Ausbruch im Mittel 0,001 Sonnenmassen verliert; die im Durchschnitt emittierte Energie entspricht der, die die Sonne in 10000 Jahren ausstrahlt. Die physikalischen Ursachen für die Novaausbrüche sind nicht sicher bekannt. Es wird vermutet, daß alle Novae Komponenten enger Doppelsternsysteme sind, eventuell sogar weiße Zwerge, die von ihrem nahen Begleiter schubweise Materie zugeführt bekommen. Der Gaseinfall löst Kernreaktionen aus, in deren Gefolge die äußeren Schichten des Sterns abgeworfen werden.

Die Gesamtzahl der in einem Jahr im Milchstraßensystem aufleuchtenden Novae dürfte etwa 50 betragen, doch wird nur ein kleiner Teil davon beobachtet.

Bei den *Supernovae* verläuft der Ausbruch mit solcher Gewalt, daß die ganze innere Struktur des Sterns beeinflußt wird. Die plötzliche Helligkeitszunahme kann mehr als 20 Größenklassen betragen; das entspricht einem Leuchtkraftanstieg um das Hundertmillionenfache!

Für wenige Tage erreicht eine Supernova die Helligkeit einer ganzen Galaxis. Die allein im optischen Spektralbereich ausgestrahlte Energie kommt der gleich, die die Sonne im Verlauf von 10 bis 100 Mill. Jahren aussendet. Im Milchstraßensystem konnten bis jetzt nur drei Supernovae beobachtet werden, die alle vor der Erfindung des Fernrohrs aufleuchteten: japanische und chinesische Chroniken berichten von einem „Gaststern" im Sternbild Taurus, der im Jahre 1054 aufleuchtete und dessen Überrest der Krebsnebel ist, 1572 entdeckte BRAHE eine Supernova im Sternbild Cassiopeia und 1604 entdeckte KEPLER eine im Sternbild Ophiuchus. Radioastronomisch konnten eine ganze Reihe von *Supernovaüberresten* ausfindig gemacht werden, die in prähistorischen Zeiten aufleuchteten und sich meist hinter dichten interstellaren Dunkelwolken verbergen. Alle anderen beobachteten Supernovae gehören extragalaktischen Sternsystemen an. Die Form der Lichtkurve ähnelt der, die bei Novae mit schnellem Helligkeitsabfall zu beobachten ist, doch ist das Helligkeitsmaximum wesentlich breiter, und die Helligkeitsschwankungen im absteigenden Ast sind seltener als bei den gewöhnlichen Novae. Als Expansionsgeschwindigkeit für die ausgeschleuderten Gasmassen wurden Werte bis zu 10000 kms^{-1} gemessen. Die Häufigkeit, mit der Supernovae in extragalaktischen Sternsystemen aufleuchten – es wird in Spiralsystemen mit 3 bis 6 Supernovae pro Jahrhundert gerechnet –, deutet darauf hin, daß es sich um massereiche Sterne handelt, die am Ende ihrer sehr schnellen Entwicklung explodieren, wobei die im Sterninneren durch Kernreaktionen erzeugten schweren Elemente in den interstellaren Raum gelangen. Neben einer expandierenden und radioastronomisch noch längere Zeit beobachtbaren Hülle ververbleibt nach dem Ausbruch noch der auf die Größe eines *Neutronensterns* geschrumpfte Kern des ehemaligen Sterns (↑ 10.3.3.9.2.). Über die vermuteten physikalischen Ursachen eines Supernovaausbruchs s. a. 10.3.3.8.1.

10.3.2. Physik der Sternatmosphären

Unter einer Sternatmosphäre wird die äußerste, lichtaussendende Schicht eines Sterns verstanden, der Teil eines Sterns also, welcher der direkten Beobachtung zugänglich ist und der der aus dem Sterninneren kommenden Strahlung in charakteristischer Weise seine physikalisch-chemischen Eigenschaften aufprägt. Gegenüber der Masse des gesamten Sterns ist die Masse seiner Atmosphäre vernachlässigbar gering; auch ist ihre Ausdehnung im allgemeinen klein gegen den Sternradius, anders wäre beispielsweise der scharf erscheinende Sonnenrand nicht zu erklären. Lediglich bei den Überriesen ist die Atmosphäre nicht mehr klein gegen den Radius.

Das Sternspektrum, aus dem auf die Struktur der Atmosphäre geschlossen wird, besteht aus einem *Kontinuum*, dem *Absorptionslinien* überlagert sind. Das Intensitätsmaximum im Kontinuum verschiebt sich mit wachsender Temperatur zu immer kürzeren Wellenlängen hin. Die Stärke einer Absorptionslinie wird sowohl durch die Menge des diese Linie hervorbringenden Elements in der Atmosphäre als auch durch die herrschenden physikalischen Bedingungen bestimmt. Die Linienabsorption hängt maßgeblich vom *Anregungs-* und *Ionisationsgrad* der betreffenden Atome ab; beide steigen aber mit zunehmender Temperatur, der Ionisationsgrad zusätzlich noch mit sinkender Dichte. Schließlich läßt die Form der Linien, die *Linienkontur*, Rückschlüsse auf die physikalischen Bedingungen zu. Die Linien werden breiter durch höheren Druck, höhere Temperatur, stark turbulente Bewegung der Materie und durch die Rotation des Sterns (↑ Abb. 10.3.1-4).

10.3.2.1. Atmosphärenschichtung. Die Theorie der Sternatmosphären erlaubt, aus dem vorliegenden Beobachtungsmaterial Rückschlüsse auf die physikalischen Verhältnisse und die chemische Zusammensetzung der Atmosphäreschichten zu ziehen. Die physikalischen Zustandsgrößen Temperatur, Dichte und Druck nehmen in Richtung auf das Sternzentrum hin zu, so daß in der Sternatmosphäre eine *Temperatur- und Druckschichtung* herrscht. Die effektive Temperatur stellt einen Temperaturmittelwert für die unteren Atmosphäreschichten, für die *Photosphäre*, die im allgemeinen klein gegen die Gesamtatmosphäre ist, dar. Sie ist gemeint, wenn von „Sternoberfläche" gesprochen wird. Die Temperatur- und Druckschichtung hängt von der chemischen Zusammensetzung der Atmosphärengase, von der Schwerebeschleunigung in der Atmosphäre und von der im Sterninneren erzeugten Energiemenge ab, die sie durchsetzt. Eine hohe Schwerebeschleunigung drückt die Sternatmosphäre zusammen, umgekehrt sind Riesensterne, die sich durch eine geringe Schwerebeschleunigung an der Oberfläche auszeichnen, von ausgedehnten atmosphärischen Hüllen geringer Dichte umgeben. Für jede Stelle der Atmosphäre gilt die *hydrostatische Gleichgewichtsbedingung*, wonach Temperatur und Dichte genau die Werte annehmen, für die der Gas- und Strahlungsdruck dem Gewicht der darüberliegenden Schichten die Waage hält. Da der Transport der im Sterninneren erzeugten Energie im wesentlichen durch Strahlung erfolgt und von den äußeren Sternschichten genau so viel Energie ausgestrahlt wird, wie aus dem Sterninneren in

sie eintritt, ist der Strahlungsstrom in der Atmosphäre konstant, und man spricht von *Strahlungsgleichgewicht*.

10.3.2.2. Strahlungstransport. Die Berechnung der Temperaturschichtung in einer Sternatmosphäre, die ja die Ursache für den nach außen gerichteten Energiestrom ist, reduziert sich mathematisch gesehen auf die Lösung der *Strahlungstransportgleichung*. Bezeichnet man mit $I_\nu(r, \vartheta)$ die orts- und richtungsabhängige Intensität des Strahlungsfelds bei der Frequenz ν, die ein Maß für die pro Quadratzentimeter, Sekunde und Raumwinkeleinheit abgegebene Energiemenge im Frequenzintervall $\nu \ldots \nu + \mathrm{d}\nu$ ist, und mit dem linearen Absorptionskoeffizienten $\varkappa_\nu$ den relativen Energieverlust, den die ν-Strahlung nach Durchlaufen einer Wegstrecke von 1 cm in der Sternmaterie erleidet, so wird bei planparalleler Schichtung die vom Flächenelement $\mathrm{d}f$ in eine um den Winkel ϑ gegen die Normale geneigte Richtung emittierte Leistung $I_\nu(r, \vartheta)\, \mathrm{d}f\, \mathrm{d}\omega$ längs der Strecke $\mathrm{d}s$ um den Betrag $\varkappa_\nu I_\nu(r, \vartheta)\, \mathrm{d}s\, \mathrm{d}f\, \mathrm{d}\omega$ verkleinert (Abb. 10.3.2-1), wobei $\mathrm{d}\omega$ das Raumwinkelelement beschreibt. Andererseits erfolgt eine Verstärkung des Strahlungsfelds durch die

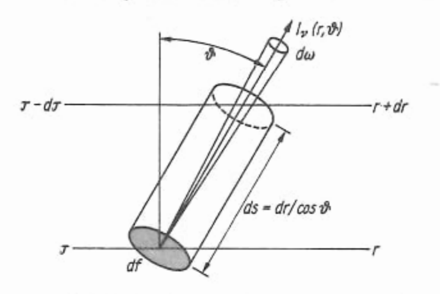


Abb. 10.3.2-1 Zur Ableitung der Grundgleichung für den Strahlungstransport

vom Volumenelement $\mathrm{d}f\,\mathrm{d}s$ reemittierte Leistung um den Betrag $\varepsilon_\nu\, \mathrm{d}f\, \mathrm{d}s\, \mathrm{d}\omega$, wobei die Größe ε_ν als Emissionskoeffizient bezeichnet wird. Er gibt die pro cm³ in die Raumwinkeleinheit reemittierte Strahlungsleistung im Frequenzintervall an. Beziehen wir uns üblicherweise nicht auf den Radius r, sondern auf die von der Oberfläche des Sterns nach innen gerichtete *optische Tiefe* gemäß $\mathrm{d}\tau_\nu = -\varkappa_\nu\, \mathrm{d}r$ und beachten den folgenden Zusammenhang $\mathrm{d}s = \mathrm{d}r/\cos\vartheta$, so kann die Gesamtbilanz des Strahlungstransports geschrieben werden:

$$(2) \quad \cos\vartheta \frac{\mathrm{d}I_\nu(\tau_\nu, \vartheta)}{\mathrm{d}\tau_\nu} = I_\nu(\tau_\nu, \vartheta) - \varepsilon_\nu/\varkappa_\nu\,.$$

Eine ähnliche Beziehung gilt auch für die über den gesamten Frequenzbereich integrierte Strahlungsintensität. Dazu ist der Absorptionskoeffizient $\varkappa_\nu$ nur durch einen geeignet definierten Mittelwert $\bar{\varkappa}$, die *Opazität*, zu ersetzen. Für jede Lösung der Strahlungstransportgleichung müssen an der Oberfläche ($\tau_\nu = 0$) der nach innen gerichtete Beitrag des Strahlungsfelds verschwinden und der frequenzintegrierte Strahlungsstrom unabhängig von der Tiefe sein.

Der materialabhängige Teil der Transportgleichung wird üblicherweise unter dem Begriff Ergiebigkeit zusammengefaßt: $S_\nu = \varepsilon_\nu/\varkappa_\nu$. Sie kann also bei bekannter chemischer Zusammensetzung, Dichte und Temperatur aus den atomaren Detailprozessen berechnet werden. Die Linienstrahlung wird zweckmäßigerweise gesondert betrachtet.

10.3.2.3. Lokales thermodynamisches Gleichgewicht. In der Theorie des Strahlungstransports sind zwei Grenzfälle von besonderem Interesse: Die *kohärente Streuung* basiert auf der elastischen Streuung der Strahlung an Atomen und Molekülen (Rayleighstreuung) oder freien Elektronen (Thomsonstreuung). Die Energie des gestreuten Photons bleibt dabei während des Streuvorgangs unverändert. Diese Art des Strahlungstransports spielt in den Atmosphären heißer Sterne eine geringe Rolle, ist aber bedeutungsvoll für die Behandlung von Planetenatmosphären und z. B. verantwortlich für die Blaufärbung des Himmels.

Die Annahme *lokalen thermodynamischen Gleichgewichts* geht davon aus, daß die absorbierte Strahlung zunächst vollständig thermalisiert wird, d. h. zur Aufrechterhaltung der Temperatur in der Sternatmosphäre dient. Nach dem Kirchhoffschen Gesetz (↑ 6.1.2.2.3.) ist dann der Emissionskoeffizient mit dem Absorptionskoeffizienten gemäß $\varepsilon_\nu = \varkappa_\nu B_\nu(T)$ verknüpft und die Ergiebigkeit ist mit der Kirchhoff-Planck-Funktion identisch:

$$(3) \quad S_\nu(\tau) = B_\nu(T(\tau)) = 2h\nu^3/c^2 \cdot 1/\{\exp[h\nu/kT] - 1\}\,.$$

Im Falle *thermodynamischen Gleichgewichts* wäre die Temperatur eine den Zustand der gesamten Atmosphäre charakterisierende Größe und unabhängig von der Tiefe.

Daß in den Sternatmosphären kein thermodynamisches Gleichgewicht herrscht, geht schon daraus hervor, daß Spektrallinien auftreten; ein im thermodynamischen Gleichgewicht befindlicher Körper strahlt ein rein kontinuierliches Spektrum aus (↑ 10.3.1.4.).

Selbst unter der Annahme lokalen thermodynamischen Gleichgewichts ist das Strahlungstransportproblem nur für Spezialfälle elementar lösbar. Als Beispiel möge die „graue" Näherung dienen. Der Absorptionskoeffizient wird hierbei durch einen geeignet gewählten Mittelwert ersetzt, und die Transportgleichung kann daraufhin sofort über die Frequenz inte-

griert werden. Die Tiefenabhängigkeit der über den gesamten Spektralbereich integrierten Intensität erlaubt nach dem Stefan-Boltzmannschen Gesetz die Berechnung der Temperaturschichtung.

Bereits diese Näherung beschreibt den gemessenen Verlauf der *Randverdunklung* für die Sonne recht gut und bestätigt damit die in die Theorie eingebrachten Voraussetzungen, wie Strahlungsgleichgewicht und lokales thermodynamisches Gleichgewicht. Die Randverdunklung ist ja sichtbarer Ausdruck für das Temperaturgefälle in der Sonnenatmosphäre und kommt dadurch zustande, daß der Beobachter in der Mitte der Sonnenscheibe in tiefere und somit heißere Schichten blickt, die entsprechend mehr Strahlung emittieren, als am Rand, wo der Sehstrahl bei unveränderter Eindringtiefe nur kühlere Atmosphärenschichten erreicht.

Im allgemeinen ist die „graue" Näherung ungeeignet, den Verlauf eines Sternkontinuums befriedigend wiederzugeben, und es muß der genaue Frequenzverlauf des Absorptionskoeffizienten berücksichtigt werden. Für das Kontinuum sind dabei die gebunden-freien Übergänge, die die Grenzkontinua erzeugen, und die frei–freien Übergänge des Wasserstoffs und seines negativen Ions, letzteres vor allem in kühlen, sonnenähnlichen Sternen, von Wichtigkeit. Der Einfluß der Absorptionslinien auf die Temperaturschichtung ist zu berücksichtigen, da die in den Linien zurückgestrahlte Energie ähnlich wie bei dem bekannten Glashauseffekt zu einer Aufheizung der tieferen atmosphärischen Schichten und damit zu einer Anhebung des Kontinuums zwischen den Linien führt. Schließlich muß in heißen Sternen und den dünnen Atmosphären der Riesen und Überriesen mit *Abweichungen vom lokalen thermodynamischen Gleichgewicht* gerechnet werden, da im Falle starker Strahlungsfelder und in Gebieten extrem geringer Gasdichte die Besetzungszahlen für die atomaren Energieniveaus nicht mehr durch Stöße allein bestimmt werden, sondern die Strahlungsanregung dieser Niveaus eine Rolle zu spielen beginnt und Abweichungen von der thermodynamischen Boltzmannverteilung auftreten.

Die Berechnung von Nichtgleichgewichtszuständen bereitet erhebliche Schwierigkeiten, weil dann die Detailprozesse der Wechselwirkung des stellaren Strahlungsfelds mit dem Atmosphärengas berücksichtigt werden müssen.

Probleme anderer Art werfen Sterne später Spektraltypen auf, deren *Wasserstoffkonvektionszonen* bis in den atmosphärischen Bereich heraufreichen, aus dem die Strahlung stammt, so daß der Energietransport teilweise konvektiver Natur ist, also durch aufsteigende heiße Gasmassen bewerkstelligt wird. Die in der Photosphäre der Sonne beobachtbaren *Granulen* dürfen als solche über die eigentliche Wasserstoffkonvektionszone hinausschießende heiße Materieballungen betrachtet werden. Auch die im Bereich einer *Absorptionslinie* ausgesandte Strahlung wird durch die Gleichung 10.3.(2) einschließlich der erwähnten Randbedingungen beschrieben. Dabei kann in vielen Fällen auf die Annahme lokalen thermodynamischen Gleichgewichts zurückgegriffen werden. Wegen der im Linienbereich besonders starken Absorptionsfähigkeit der Sternmaterie, stammt die Linienstrahlung im allgemeinen aus höheren und somit kühleren atmosphärischen Schichten als das die Linie einbettende Kontinuum. Die im Linienkern beobachtete Restintensität ist ein Maß für die Temperatur der die Linie erzeugenden Schichten.

10.3.2.4. Bestimmung der chemischen Häufigkeiten. Bereits das Auftreten einer bestimmten Absorptionslinie erlaubt eine rein qualitative Aussage über die Präsenz des die Linie erzeugenden chemischen Elements in der Atmosphäre. Eine quantitative Bestimmung der chemischen Zusammensetzung läßt sich erst durchführen, wenn die physikalischen Verhältnisse, d. h. Temperatur- und Druckschichtung, bekannt sind; diese aber ergeben sich aus der Häufigkeit der Elemente, da diese wesentlich die Wechselwirkung zwischen Strahlung und Materie mitbestimmen. Eine geschlossene Lösung des Problems kann es deshalb nicht geben. Man muß auf Näherungsverfahren zurückgreifen.

In einer ersten *Grobanalyse* geht man von einer isothermen und isobaren Sternatmosphäre aus, für die man Temperatur- und Druckmittelwert aus dem beobachteten Spektrum ableitet. Aus der Stärke der Spektrallinien ergeben sich dann Mittelwerte für die Häufigkeit der verschiedenen Elemente, die schon recht gut mit der tatsächlichen Häufigkeit übereinstimmen. Im nächsten Schritt des Näherungsverfahrens, der *Feinanalyse*, berechnet man mit dieser chemischen Häufigkeit und einem Mittelwert für die Schwerebeschleunigung den Temperatur- und Dichteverlauf unter der Bedingung konstanten Strahlungsstroms. Verbleibende Abweichungen der aus diesem verbesserten Atmosphärenmodell folgenden theoretischen Linienstärken und -konturen gegenüber dem beobachteten Linienspektrum können dann durch kleine Korrekturen der Elementehäufigkeiten und der Schwerebeschleunigung weitgehend beseitigt werden. Mit diesen neuen Werten wird das nächste Näherungsmodell für die Atmosphäre bestimmt. Das Verfahren wird fortgesetzt, bis befriedigende Übereinstimmung zwischen den beobachteten und den berechneten spektralen Daten erreicht ist. Die bisherigen Analysen

ergaben, daß die chemische Zusammensetzung der meisten untersuchten Sternatmosphären sehr ähnlich ist. Aus Untersuchungen von Sternatmosphären, vor allem der Sonnenatmosphäre, aber auch von Gasnebeln und Meteoriten erhält man die mittlere kosmische Häufigkeit der Elemente (Tab. 10.3.2-2).

Tab. 10.3.2-2 Mittlere kosmische Elementenhäufigkeit bezogen auf eine Siliciumhäufigkeit von 10^6 Atomen

Element	Häufigkeit	Element	Häufigkeit
H	$3{,}2 \cdot 10^{10}$	S	$3{,}8 \cdot 10^5$
He	$4{,}1 \cdot 10^9$	Cl	8800
Li	100	Ar	$1{,}5 \cdot 10^5$
Be	20	K	3160
C	$1{,}1 \cdot 10^7$	Ca	$4{,}9 \cdot 10^4$
N	$3{,}0 \cdot 10^6$	Ti	2400
O	$3{,}1 \cdot 10^7$	Cr	7800
F	1600	Mn	6850
Ne	$8{,}6 \cdot 10^6$	Fe	$1{,}0 \cdot 10^4$
Na	$4{,}4 \cdot 10^4$	Ni	$2{,}7 \cdot 10^4$
Mg	$9{,}1 \cdot 10^5$	Cu	212
Al	$9{,}5 \cdot 10^4$	Zn	490
Si	$1{,}0 \cdot 10^6$	Zr	54
P	$1{,}0 \cdot 10^4$		

10.3.3. Innerer Aufbau und Entwicklung der Sterne

Da das Sterninnere mit herkömmlichen Instrumenten nicht beobachtbar ist, müssen alle Aussagen über seine physikalische Struktur und seine chemische Zusammensetzung aus theoretischen Überlegungen erschlossen werden. Die von der Theorie des inneren Aufbaus und der Entwicklung der Sterne erarbeiteten Modellvorstellungen ermöglichen ein tieferes Verständnis der beobachteten Hertzsprung-Russell-Diagramme (↑ 10.3.1.9.). Außerdem vermag die Theorie der Sternentwicklung die Frage nach der Herkunft der schweren Elemente im Kosmos zu beantworten, die entgegen älteren Ansichten nicht im heißen Feuerballstadium des Universums entstanden sein können, sondern erst als Folge von Kernumwandlungen im Inneren massereicher Sterne synthetisiert wurden.

10.3.3.1. Mechanisches Gleichgewicht. Nach Beobachtungen ändern sich die Zustandsgrößen der Sterne wie Radius, Leuchtkraft, effektive Temperatur u. a. nicht merklich. Deshalb kann sich auch ihre innere Struktur, von der die Zustandsgrößen abhängen, höchstens in langen Zeiträumen ändern; selbst bei den Pulsationsveränderlichen kann der mittlere Zustand der Sterne, um den der Stern mehr oder minder regelmäßig schwingt, weitgehend als konstant angesehen werden. Diese Beobachtungsergebnisse beziehen sich nur auf einen im Vergleich zur gesamten Existenzzeit eines Sterns außerordentlich kurzen Zeitabschnitt, für die Sonne aber kann gezeigt werden, daß sie mindestens während der letzten 10^9 Jahre mit praktisch der gleichen Intensität gestrahlt haben muß wie im Augenblick. Bei geologischen Untersuchungen wurden fossile Algen gefunden, die mindestens 10^9 Jahre alt sind und die für ihre Existenz etwa die gleiche mittlere Temperatur brauchten, wie sie augenblicklich auf der Erde herrscht. Die maximalen Abweichungen von der jetzigen mittleren Temperatur dürften höchstens 20 K betragen haben. Radius, Leuchtkraft und effektive Temperatur der Sonne können sich danach in diesem Zeitraum nur in äußerst geringem Maße geändert haben. Dies gilt dann auch für ihre innere Struktur und, da die Sonne ein ganz normaler Stern ist, für die Mehrzahl aller anderen Sterne.

Da die überwiegende Mehrheit der Sterne nach den Beobachtungen weder stark rotieren noch starke Magnetfelder oder sehr nahe Begleiter haben, müssen sie kugelsymmetrisch aufgebaut sein, so daß man den physikalischen Zustand und die chemische Zusammensetzung nur längs eines Radiusvektors zu kennen braucht. Statt eines dreidimensionalen Problems muß nur ein lineares gelöst werden.

Die innere Stabilität eines Sterns erfordert, daß zu jedem Zeitpunkt und in jedem Volumenelement der Druck dem Gewicht der über diesem Gebiet lagernden Sternmaterie die Waage hält; er wird aus diesem mechanischen oder hydrostatischen Gleichgewicht bestimmt. Von seinen beiden Komponenten ist der Strahlungsdruck meist klein im Vergleich zum Gasdruck.

10.3.3.2. Energetisches Gleichgewicht. Die innere Stabilität der Sterne erfordert weiterhin, daß genau so viel Energie ausgestrahlt wird, wie im Inneren erzeugt wird. Es kann daher im Inneren weder Gebiete geben, in denen dauernd Energie gespeichert wird, noch solche, die durch starke Ausstrahlung dauernd kühler werden, da sonst eine dauernde Temperatur- und damit Druckänderung eine Expansion oder eine Kontraktion und damit eine Änderung der inneren Struktur des Sterns zur Folge hätte.

10.3.3.3. Zustandsgleichungen. Der physikalische Zustand der Sternmaterie läßt sich im allgemeinen mit der Gleichung für ideale Gase beschreiben, z. B. sogar in den zentrumsnahen Gebieten der Sonne, in denen die Dichte mehr als 100 gcm^{-3} beträgt. Die Gültigkeit der idealen Gasgleichung wird durch eine hohe Temperatur gesichert, in der Sonnenmitte z. B. rund $14 \cdot 10^6$ K, bei der alle Atome praktisch vollständig ionisiert sind und fast sämtliche Elektronen verloren haben. Das Eigenvolumen der

Elektronen und der Atomkerne ist wesentlich kleiner als das von unionisierten Atomen und kann gegenüber dem Raum, der einem einzelnen freien Teilchen zur Verfügung steht, vernachlässigt werden. Erst bei viel größeren Dichten, z. B. von 10^5 gcm^{-3} bei den Weißen Zwergen, verliert diese Gleichung ihre Gültigkeit und muß durch eine andere Zustandsgleichung für *entartete* oder *überdichte Materie* ersetzt werden, nach der z. B. der Druck nur von der Dichte und nicht auch von der Temperatur abhängt.

10.3.3.4. Energiefreisetzungsprozesse. Die im Stern erzeugte Energie wird fast ausschließlich bei Kernreaktionen frei, durch die aus leichten Atomkernen schwerere aufgebaut werden. Die dem Massendefekt nach dem Einsteinschen Äquivalenzprinzip $E = mc^2$ äquivalente Energie wird ausgestrahlt. Die wichtigsten Prozesse sind dabei die Proton-Proton-Kette, der Kohlenstoff-Stickstoff-Zyklus und die Heliumreaktion.

Neben den Kernprozessen kommt für die Energieerzeugung noch die Kontraktion des gesamten Sterns oder seiner Teile in Frage. Die dabei freigesetzte potentielle Energie wird in thermische umgewandelt, die zum Teil die Sternmaterie aufheizt und zum anderen Teil ausgestrahlt wird. Bei Kontraktion erhöht sich immer die Temperatur; bei einer Expansion, die in manchen Phasen der Sternentwicklung in den äußeren Gebieten vorkommen kann, wird die Temperatur erniedrigt.

10.3.3.4.1. *Proton-Proton-Kette.* Der wirksame Ablauf dieser Kette besteht aus drei Schritten (Abb. 10.3.3-1); aus zwei Protonen H-1 bildet sich ein Deuteron H-2 unter Freigabe von einem Positron e^+ und einem Neutrino ν. Das Positron reagiert fast sofort mit einem Elektron unter Paarvernichtung und Ausstrahlung zweier γ-Quanten; das Neutrino verläßt den Stern ungehindert. Beim Stoß des Deuterons mit einem weiteren Proton wird ein Helium-3-Kern unter Aussendung eines γ-Quants gebildet. Zwei solcher Helium-3-Kerne bilden beim Stoß einen Helium-4-Kern unter Freigabe von zwei Protonen. Insgesamt gehen sechs Protonen (Wasserstoffkerne) in diesen Prozeß ein; vier vereinigen sich unter Abgabe einer Energie von insgesamt $4{,}2 \cdot 10^{-12}$ J zu einem Helium-4-Kern, und zwei stehen für weitere Reaktionen wieder zur Verfügung. Die Häufigkeit, mit der der Prozeß in einem Stern abläuft, hängt von der herrschenden Dichte und von der Temperatur ab. Sie wächst mit dem Quadrat der Dichte und mit der 4. bis 6. Potenz der Temperatur. Zur Erzeugung merklicher Energiemengen sind Temperaturen von einigen Millionen Kelvin notwendig.

Abb. 10.3.3-1 Proton-Proton-Kette

10.3.3.4.2. *Kohlenstoff-Stickstoff-Zyklus.* In ihm entsteht der Helium-4-Kern dadurch aus vier Protonen, daß diese an den Kohlenstoff-12-Kern oder seine Folgeprodukte angelagert werden. Einige dieser Kerne sind instabil und wandeln sich in andere um, insbesondere wird am Schluß ein C-12-Kern wieder frei, so daß Kohlenstoff im Zyklus die Rolle eines Katalysators spielt (Abb. 10.3.3-2). Die insgesamt dabei erzeugte Energie beträgt $4 \cdot 10^{-12}$ J. Zum Ablauf ist eine Temperatur von über $12 \cdot 10^6$ K notwendig, die mithin größer ist als die für die Proton-Proton-Kette. Da die Energieerzeugung durch den Kohlenstoff-Stickstoff-Zyklus mit der 16. bis 20. Potenz der Temperatur steigt, ist sie über rund $16 \cdot 10^6$ K ergiebiger als die durch die Proton-Proton-Kette. Die Gebiete, in denen die Energieerzeugung in einem Stern stattfindet, sind wegen der starken Temperaturabhängigkeit nahe um den Mittelpunkt konzentriert.

Abb. 10.3.3-2 Kohlenstoff-Stickstoff-Zyklus

10.3.3.4.3. *Heliumreaktion.* Herrschen im Sterninneren Temperaturen von mindestens 8 bis $10 \cdot 10^7$ K, so können aus Heliumkernen über den Beryllium-8-Kern Kohlenstoffkerne etwa nach dem Schema He-4 + He-4 → Be-8 + He-4 → C-12 + γ gebildet werden. Der Energiegewinn beträgt dabei $1{,}73 \cdot 10^{-12}$ J. Über Prozesse, noch schwerere Kerne zu bilden, und über die dabei frei werdende Energie sind die Kenntnisse verhältnismäßig unsicher. Aus C-12 kann z. B. Mg-24 unter Aussendung eines γ-Quants oder Na-23 unter Aussendung eines Protons entstehen. Diese Prozesse liefern aber erst dann genügend Energie, wenn die Tem-

peratur mindestens rund $6 \cdot 10^6$ K beträgt (Abb. 10.3.3-3).

10.3.3.5. Energietransport. Die im Sterninneren erzeugte Energie wird zum größten Teil als Strahlung nach außen transportiert. Die bei der Kernumwandlung und durch Paarvernichtung entstandenen Energiequanten können innerhalb des Sterns nur sehr kurze Wegstrecken zurücklegen; es läßt sich zeigen, daß schon eine Materieschicht von wenigen Zentimetern Dicke ausreicht, alle Energiequanten zu absorbieren. Als *Absorptionsprozeß* wirken gebunden–freie und frei–freie Übergänge der Elektronen sowie die Streuung der Energiequanten an den freien Elektronen. Die Energiequanten werden nach der Absorption nach allen Richtungen wieder ausgestrahlt.

In Sterngebieten, in denen ein starkes Temperaturgefälle von innen nach außen herrscht, kann auch *Konvektion* den Energietransport übernehmen. Wenn sie einsetzt, ist sie so wirksam, daß der Strahlungstransport nicht mehr ins Gewicht fällt. Bei der Konvektion steigen heiße Gasmassen im Stern auf, während kühle absinken. Der Energietransport ist dann mit einem Materietransport verknüpft.

10.3.3.6. Eindeutigkeitssatz, Sternmodelle. Aus den Untersuchungen der Gleichgewichtsbedingungen, der Energieerzeugung und des Energietransports geht unter Berücksichtigung der Zustandsgleichungen hervor, daß der innere Aufbau eines Sterns allein durch die *Sternmasse* und die *chemische Zusammensetzung* bestimmt ist. Alle anderen Größen, z. B. die der Beobachtung zugänglichen Zustandsgrößen wie Radius, Leuchtkraft und effektive Temperatur sind damit auch durch diese Parameter festgelegt. Gibt man Masse und chemische Zusammensetzung vor, so läßt sich der innere physikalische Zustand eines Sterns eindeutig berechnen. Die Zusammenfassung dieser numerischen Daten wird als *Sternmodell* bezeichnet.

Soll für einen Stern mit bekannter Masse der innere Aufbau bestimmt werden, so müssen mit unterschiedlichen Ansätzen für die chemische Zusammensetzung verschiedene Sternmodelle gerechnet und die sich ergebenden äußeren Zustandsgrößen mit den tatsächlich beobachteten verglichen werden. Es zeigt sich, daß für die *Hauptreihensterne* die beste Übereinstimmung zu erreichen ist, wenn für den gesamten Stern eine gleichmäßige Zusammensetzung aus 60 bis 90% Wasserstoff, 1 bis 3% Elementen schwerer als Helium und dem Rest aus Helium angenommen wird. Aus den Modellrechnungen ergibt sich weiterhin, daß die *Zentraltemperatur* massereicher Sterne höher als die massearmer ist, die *Zentraldichte* umgekehrt aber niedriger. Ein Stern mit 7 Sonnenmassen hat bei dieser chemischen Zusammensetzung eine Zentraltemperatur von rund $29 \cdot 10^6$ K und eine Zentraldichte von 11 $\mathrm{gcm^{-3}}$, die Sonne hingegen von rund $14 \cdot 10^6$ K und 110 $\mathrm{gcm^{-3}}$ (Abb. 10.3.3-4). In den Hauptreihensternen mit Massen größer als etwa 2 Sonnenmassen wird die Energie hauptsächlich durch den *Kohlenstoff-Stickstoff-Zyklus* erzeugt. Wegen der starken Konzentration der Energiequellen zum Zentrum und wegen des sich daher einstellenden starken Temperaturgefälles wird die Energie in den zentrumsnahen Gebieten dieser Sterne durch *Konvektion* transportiert. Bei den Sternen mit Massen kleiner oder gleich einer Sonnenmasse ist die *Proton-Proton-Kette* für

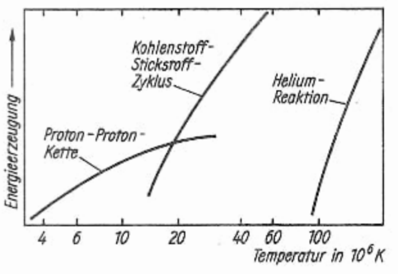


Abb. 10.3.3-3 Abhängigkeit der verschiedenen Energieerzeugungsprozesse von der Temperatur

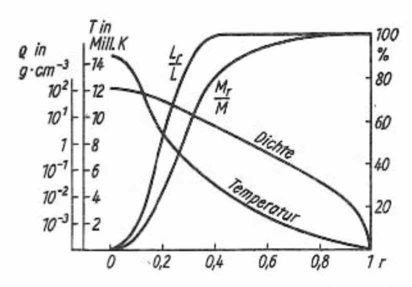


Abb. 10.3.3-4 Temperatur- und Dichteverlauf in der Sonne sowie der Verlauf der relativen Masse M_r/M, die sich innerhalb einer Kugel mit dem relativen Radius r (in Einheiten des Sonnenradius) befindet, und der Verlauf der relativen Leuchtkraft L_r/L

die Energieerzeugung wesentlich. Konvektion tritt bei ihnen nur in den Außenbezirken auf, in denen die Temperatur so niedrig ist, daß der Wasserstoff nur teilweise ionisiert ist. Es herrscht dann eine starke Absorption, die ein großes Temperaturgefälle und damit Konvektion verursacht.

Messungen des *solaren Neutrinoflusses* haben in neuerer Zeit Zweifel an der Richtigkeit der berechneten Sonnenmodelle aufkommen lassen. Die nachgewiesenen Neutrinos stammen, da die Sonne für die Neutrinostrahlung durchsichtig

ist, unmittelbar aus der Wasserstoffbrennzone, in der sie als Nebenprodukt bei der Energieerzeugung anfallen (↑ 10.3.3.4.). Die Neutrinoproduktionsraten und das Energiespektrum der Neutrinostrahlung hängen empfindlich von der Art der Energieerzeugung und somit von der im Sonneninneren herrschenden Temperatur ab. Nachgewiesen werden die Sonnenneutrinos gemäß der von PONTECORVO und ALVAREZ vorgeschlagenen Reaktion Cl-37 + $\bar{\nu}_e \rightarrow$ Ar-37 + e^- − 0,814 MeV. Die entstandenen Argonatome sind radioaktiv und zerfallen mit einer Halbwertszeit von 35 Tagen wieder zu Chlor. Die Einfangwahrscheinlichkeit für Neutrinos ist allerdings überaus klein. Die Experimentatoren stehen vor der Aufgabe, in einem mit 610 t flüssigem Perchlorethylen (C_2Cl_4) gefüllten und sorgfältig gegen äußere Störeinflüsse abgeschirmten Tank 35 radioaktive Argonatome zahlenmäßig nachzuweisen, da theoretisch mit einer einzigen Einfangreaktion pro Tag gerechnet wurde. Tatsächlich werden weniger Zerfallsereignisse als erwartet registriert. Die aufgetretene Diskrepanz kann möglicherweise auch auf eine fehlerhafte Neutrinotheorie zurückzuführen sein. Eine befriedigende Erklärung bietet sich durch Einbeziehung von drei Neutrinoarten an die sich untereinander ins Gleichgewicht setzen.

10.3.3.7. Sternentstehung. Während sich über den inneren Aufbau der Hauptreihensterne und auch über die nachfolgenden Entwicklungsetappen relativ gesicherte Angaben machen lassen, wird der Prozeß der Sternentstehung bis heute noch nicht in der wünschenswerten Ausführlichkeit verstanden.

Daß auch heute noch Sterne entstehen, geht z. B. daraus hervor, daß sich massereiche Hauptreihensterne der Spektralklasse 0 nur die verhältnismäßig kurze Zeit von einigen Millionen Jahren in diesem Zustand aufhalten und deshalb keinesfalls älter sein können. Zum anderen beobachtet man, daß sich frühe, heiße Sterne immer in räumlicher Nähe *interstellarer Materie* befinden, aus der sie sich offensichtlich gebildet haben.

Die Raumgeschwindigkeit dieser Sterne ist nämlich im allgemeinen viel zu gering, als daß sie sich weit von ihrem Ursprungsort entfernt haben können. Die den Raum zwischen den Sternen ausfüllende gas- und staubförmige Materie kommt in wolkenartigen Verdichtungen vor, und es scheint so, daß zumindest die größten interstellaren Wolken mit einer Masse von einigen 10^4 Sonnenmassen kurz vor dem Gravitationskollaps stehen. Bereits ein geringfügiger äußerer Druckanstieg läßt dann eine solche Wolke unter der Wirkung ihrer Eigenschwere zusammenbrechen.

Auf Fotografien benachbarter Spiralgalaxien erkennt man deutlich, daß die jungen, heißen Sterne längs der Spiralarme perlschnurartig aufgereiht sind (Tafel 61). Da sich die Spiralarme durch eine hohe Konzentration von neutralem Wasserstoffgas auszeichnen (↑ 10.3.4.1.), dem „Baumaterial" für die Sterne, nimmt man an, daß eine großräumige, mit dem Spiralarmphänomen verbundene Kompressionswelle im interstellaren Gas den Kollaps großer Gas-Staub-Wolken und damit die Bildung neuer Sterne in den Spiralarmen auslöst (↑ 10.3.5.3.).

Eine einmal instabil gewordene Wolke fällt unaufhaltsam in sich zusammen. Bei einer angenommenen Anfangsdichte von nur 100 Wasserstoffatomen pro cm^3 nimmt der Kontraktionsprozeß lediglich einige 10^7 Jahre in Anspruch. Die Bildung von Protosternen geht also vergleichsweise schnell vor sich. Bevor allerdings Einzelsterne entstehen können, muß die sich immer mehr verdichtende Wolke eine *Fragmentierungsphase* durchmachen, sie zerbricht in viele, nunmehr für sich kontrahierende Teile. Dieser für die Deutung des beobachteten Massenspektrums der Sterne wesentliche Prozeß der Fragmentierung wird bisher nur in den Grundzügen verstanden.

Die weitere Entwicklung eines aus der Ursprungswolke durch Fragmentierung hervorgegangenen Bruchstücks von beispielsweise einer Sonnenmasse läßt sich rechnerisch verfolgen. Bereits nach 400000 Jahren hat sich im Zentrum ein optisch dichter Kern herausgebildet, der sich nur noch langsam verdichtet aber stark aufheizt, da er die bei seiner weiteren Kontraktion freiwerdende Gravitationsenergie nicht mehr ungehindert in Form von Wärmestrahlung abgeben kann. Seine weitere Entwicklung kann durch eine Folge von *hydrostatischen Gleichgewichtszuständen* hinreichend genau beschrieben werden. Das von außen auf den dichten Kern – einige 10^{-12} gcm^{-3} – niederregnende Hüllenmaterial heizt den Kern beträchtlich auf, außerdem vergrößert sich seine Masse. Reicht die im Inneren erreichte Temperatur zur Dissoziation der Wasserstoffmoleküle aus, wird der Kern erneut instabil, da die zur Dissoziation benötigte Wärmemenge dem Gas entzogen wird und der Gasdruck dadurch schlagartig sinkt. Der *Protostern* schrumpft auf etwa den heutigen Sonnendurchmesser zusammen. Noch in der Folgezeit ist der entstandene Protostern von einer optisch dichten Hülle umgeben, so daß er sich einem außenstehenden Beobachter nur durch eine ungewöhnlich hohe *Infrarotstrahlung* bemerkbar macht, die durch Thermalisation des Sternenlichts an den Staubteilchen der Hülle entsteht. Tatsächlich sind aus einigen Gebieten unserer Milchstraße, in denen der Sternentstehungsprozeß noch in vollem Gange zu sein scheint, einige starke Infrarotstrahlungsquellen bekannt, die mit diesem Stadium der Protosternentwicklung in

Verbindung gebracht werden. Erst nach dem „Aufklaren" der äußeren Hülle wird der neu entstandene Stern, dessen Temperatur im Inneren inzwischen so weit angestiegen ist, daß die Wasserstoffusion als Energiequelle erschlossen ist, schließlich selbst sichtbar. Er hat dann den seiner Masse entsprechenden Platz auf der Hauptreihe eingenommen. Möglicherweise befinden sich die *T-Tauri-Sterne*, wegen ihrer räumlichen Nähe zur interstellaren Materie auch als *Nebelveränderliche* bezeichnet, in der Schlußphase ihrer kontraktiven Entwicklung, kurz vor dem Einsetzen des „Wasserstoffbrennens".

10.3.3.8. Sternentwicklung. Da nach dem Eindeutigkeitssatz der innere Aufbau der Sterne allein durch Masse und chemische Zusammensetzung bestimmt wird, die chemische Zusammensetzung sich aber wegen der Kernprozesse ändert, ist die innere Struktur eines Sterns notwendigerweise zeitabhängig. Die Energieerzeugungsrate und damit die Umwandlungsrate eines Elements in ein anderes hängen von Dichte und Temperatur ab und lassen sich für jeden Punkt des Sterninneren bestimmen. Aus der angenommenen chemischen Zusammensetzung der Sternmaterie zu einem Ausgangszeitpunkt kann sie danach auch für spätere Zeitpunkte berechnet werden.

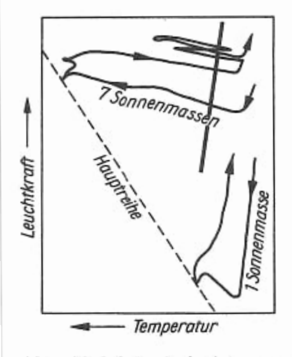


Abb. 10.3.3-5 Beispiele von Entwicklungswegen im Hertzsprung-Russell-Diagramm; *dicke Linie* Bereich der δ-Cephei-Sterne

Die Änderung der chemischen Zusammensetzung erfolgt im tiefen Inneren des Sterns wegen der dort ablaufenden Kernprozesse; in den äußeren Schichten bleibt sie praktisch unverändert. Da die Sternmasse im allgemeinen während dieser Zeit konstant bleibt, läßt sich aus einem Ausgangsmodell eine Folge von anderen Modellen herleiten und damit die Entwicklung eines Sterns berechnen. Mit dem inneren Aufbau des Sterns ändern sich aber auch die beobachtbaren Zustandsgrößen, so daß der Stern durch seine zeitliche Entwicklung verschiedene Orte im Hertzsprung-Russell-Diagramm einnimmt. Er durchläuft dabei einen Entwicklungsweg (Abb. 10.3.3-5). Die Bedingung konstanter Masse ist für alle Sterne, für die im allgemeinen Sternmodelle gerechnet werden, in sehr guter Näherung erfüllt.

Als Ausgangszustand für Entwicklungsrechnungen wird im allgemeinen der Hauptreihenzustand angenommen, bei dem einheitliche chemische Zusammensetzung im gesamten Stern herrscht und die Kernprozesse gerade erst eingesetzt haben.

10.3.3.8.1. *Entwicklung massereicher Sterne.* Wie in großen Zügen die Entwicklung verläuft, sei an einem Stern mit $M_r = 7M_0$ (Sonnenmassen) nach Hofmeister, Kippenhahn und Weigert dargestellt. Der Stern soll zur Population I gehören und im Hauptreihenzustand die folgende chemische Zusammensetzung haben: 60,2% Wasserstoff, 35,4% Helium und 4,4% schwerere Elemente entsprechend der mittleren kosmischen Häufigkeit (↑ Tab. 10.3.2-2). Der Anfangsradius dieses Sterns beträgt $2{,}4 \cdot 10^{11}$ cm = 3,5 Sonnenradien, die Zentraltemperatur $28{,}8 \cdot 10^6$ K, die Zentraldichte 11 gcm^{-3}, die Leuchtkraft $8{,}2 \cdot 10^{36}$ erg = $8{,}2 \cdot 10^{29}$ J = $2{,}1 \cdot 10^3$ Sonnenleuchtkräfte, die effektive Temperatur $21 \cdot 10^3$ K. Die Energieerzeugung erfolgt durch den Kohlenstoff-Stickstoff-Zyklus in den innersten Teilen des Sterns, die rund 10% der Gesamtmasse enthalten. Die inneren rund 27% der Masse werden infolge der dort herrschenden Konvektion durchmischt, d. h., das bei den Kernprozessen gebildete Helium wird gleichmäßig in diesem Teil des Sterns verteilt. Weiter nach außen gelangt es nicht. Während der folgenden rund $20 \cdot 10^6$ Jahre ändert sich an dem Zustand wenig, der Stern expandiert etwas, die Leuchtkraft steigt an, obwohl die effektive Temperatur sinkt. Im Hertzsprung-Russell-Diagramm wandert der Stern von der Hauptreihe weg etwas nach rechts. Sobald im Zentrum die Wasserstoffhäufigkeit merklich geringer geworden ist, beschleunigt sich die Strukturänderung. Nach etwa $26 \cdot 10^6$ Jahren ist aller Wasserstoff im Zentrum aufgebraucht, die Energieerzeugung hört damit dort zunächst auf. Die Energiequellen haben sich nach außen verschoben; der Energiebedarf wird da gedeckt, wo noch Wasserstoff zur Verfügung steht, nachdem dort infolge von Kontraktion der inneren Gebiete die Temperatur so weit angestiegen ist, daß Kernprozesse ablaufen können. Noch bevor der Kern „leergebrannt" ist, hat sich ein *Schalenquellenmodell* herausgebildet. Schließlich erfolgt das „Wasserstoffbrennen" ganz in der Schale, die den Kern umgibt, der zu 95,6% aus Helium besteht. Während der Heliumkern infolge Kontraktion immer heißer wird, schiebt sich die *„wasserstoffbrennende" Schale* immer weiter

nach außen. Nach rund $26{,}5 \cdot 10^6$ Jahren ist die Zentraltemperatur so hoch, daß „*Heliumbrennen*" einsetzen kann und dadurch Kohlenstoff erzeugt wird. Während der Kontraktion des Heliumkerns expandieren die äußeren Sternschichten, die rund 80% der Masse enthalten (Abb. 10.3.3-6); der Stern hat sich in einen

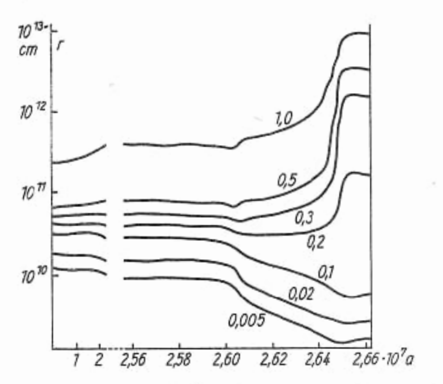


Abb. 10.3.3-6 Radiusänderungen einiger Massenschalen in einem Stern von 7 Sonnenmassen; die Zahlen an den Kurven geben die relative Masse an, die sich innerhalb der entsprechenden Massenschale befindet; während die inneren Teile (10% der Sternmasse) kontrahieren, expandieren die übrigen Gebiete; man beachte die Änderung in der Zeitskala und die logarithmische Ordinatenachse (nach HOFMEISTER, KIPPENHAHN, WEIGERT)

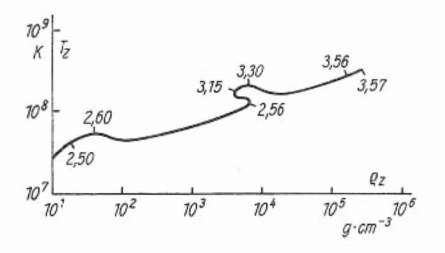


Abb. 10.3.3-7 Entwicklung von Zentraldichte ϱ_z und -temperatur T_z eines Sterns von 7 Sonnenmassen; die Zahlen geben das Alter in 10^7 Jahren nach Verlassen der Hauptreihe an (nach HOFMEISTER, KIPPENHAHN, WEIGERT)

Roten Riesen verwandelt, dessen Radius etwa 140 Sonnenradien beträgt. Die effektive Temperatur ist auf 6800 K gesunken, die Leuchtkraft hat sich daher kaum geändert. Ist der Heliumvorrat des Kerns erschöpft, verschiebt sich das Heliumbrennen in eine Schale nach außen. Damit ist folgende Struktur erreicht: zentrale *Kohlenstoffkugel*, darum heliumbrennende Schale, Heliumschale ohne Energieerzeugung, wasserstoffbrennende Schale, ungestörte Anfangszusammensetzung. Die wasserstoffbrennende Schale stirbt nach etwa $35{,}7 \cdot 10^6$ Jahren ab. Die Zentraltemperatur ist dann auf $31{,}3 \cdot 10^8$ K, und die Zentraldichte ist auf $2{,}3 \cdot 10^5$ gcm^{-3} gestiegen (Abb. 10.3.3-7). Die Dichte ist so hoch, daß nicht mehr mit einem idealen Gas gerechnet werden kann, sondern Entartung berücksichtigt werden muß.

Während sich der Stern zu einem Roten Riesen entwickelt, durchquert er auch mehrmals einen Bereich des Hertzsprung-Russell-Diagramms, der von den *Pulsationsveränderlichen*, den σ-Cephei-Sternen, besetzt ist. Die Entwicklungsrechnungen bestätigen, daß genau in dem beobachteten Instabilitätsstreifen (↑ Abb. 10.3.3-5) auch die gerechneten Sternmodelle zu Pulsationen neigen. Da die Leuchtkräfte in diesem Stadium bekannt sind und sich die Grundfrequenz für radiale Pulsationen leicht für ein vorgegebenes Sternmodell berechen läßt, kann eine theoretische *Perioden-Helligkeits-Beziehung* abgeleitet werden, die den beobachteten Kurvenverlauf gut wiedergibt (Abb. 10.3.3-8).

In den folgenden Entwicklungsphasen eines massereichen Sterns muß der Energieverlust durch Neutrinos berücksichtigt werden. Die den Stern ungehindert passierenden Neutrinos zögern die Zündung der zentralen Kohlenstoff-Sauerstoff-Kugel solange hinaus, bis Zentraldichten von 10^9 gcm^{-3} und mehr erreicht sind. Für Sterne unter 10 Sonnenmassen setzt das „Kohlenstoffbrennen" in der hochgradig entarteten Materie schlagartig ein, weil im Entartungsfall der Gasdruck nahezu unabhängig von der Temperatur ist und die einsetzenden Kernreaktionen, die ihrerseits empfindlich von der Temperatur abhängen, eine Temperaturlawine auslösen, die erst dann gestoppt wird, wenn schließlich die Entartung aufgehoben wird und sich das Kerngebiet durch Expansion abkühlt.

Es liegt nahe, in dieser *Kohlenstoffdetonation* die physikalische Ursache für die gelegentlich zu beobachtenden Sternexplosionen, die *Supernovae*, zu sehen. Gegen diese Deutung spricht eigentlich nur, daß ein in dieses kritische Stadium gelangter Stern vollständig auseinandergerissen wird, während zumindest von zwei Supernovaausbrüchen kompakte Restkörper (Krebsnebel- und Velanebelpulsar), wahrscheinlich *Neutronensterne*, bekannt sind.

Für Sterne mit mehr als 10 Sonnenmassen ist ein anderer Supernovamechanismus denkbar, der auf der thermischen Spaltung der Atomkerne des Eisens, einer stark endothermen Reaktion, beruht. Durch sukzessiven Einfang von α-Teilchen können nämlich in dem Maße, wie die Temperatur im Sterninneren ansteigt, alle chemischen Elemente bis hinauf zum Eisen, dem

Element mit der größten Bindungsenergie pro Nukleon (↑ 3.3.2.3.1.), aufgebaut werden. Bei Temperaturen von $6 \cdot 10^9$ K und einer Dichte von 10^8 gcm^{-3} setzt die *Dissoziation des Eisens* ein, Eisenkerne zerfallen spontan in Helium. Die dazu benötigte Energie wird dem Sterninneren entzogen, das dadurch dynamisch in-

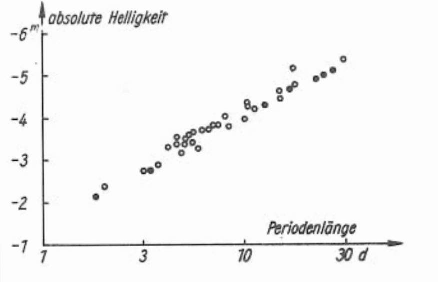


Abb. 10.3.3-8 Perioden-Helligkeits-Beziehung der δ-Cephei-Sterne; *Kreise* Beobachtung, *Punkte* Theorie (nach HOFMEISTER)

stabil wird und unter der Wirkung seines Eigengewichts in sich zusammenstürzt. Während sich im Zentrum ein extrem verdichteter Kern herausbildet, der in einem vergleichsweise winzigen Gebiet von nur wenigen Kilometern Durchmesser eine Masse von der Größe der Sonne in sich vereint, expandieren die durch eine nach außen laufende Stoßwelle aufgeheizten äußeren Schichten des Sterns. Einsetzende Kernprozesse bewirken, daß sich die äußere Hülle immer mehr aufbläht und schließlich in den interstellaren Raum abgeblasen wird. Der beschriebene Vorgang erinnert stark an die beobachteten Masseauswürfe bei Supernovaexplosionen (s. a. 10.3.1.10.). Im Gefolge des Ausbruchs werden dem interstellaren Gas gewaltige Mengen der im Sterninneren erzeugten schweren Elemente zugeführt. Dieser Anreicherungsprozeß ist für das Verständnis der empirisch bestimmten „Metallhäufigkeiten" in den Sternatmosphären von Sternen unterschiedlichen Alters wesentlich. Junge Sterne, die sich aus dem bereits mit schweren Elementen angereicherten interstellaren Gas gebildet haben, weisen im Mittel einen höheren „Metallgehalt" auf als alte Sterne.

10.3.3.8.2. *Entwicklung massearmer Sterne.* Sie erfolgt wesentlich langsamer als ein 7-Sonnenmassen-Stern, unsere Sonne hat sich beispielsweise in den 4,5 Mrd. Jahren ihres Bestehens noch nicht von der Hauptreihe entfernt, es gibt aber auch bemerkenswerte qualitative Unterschiede. In massearmen Sternen wird ein Großteil der Energie durch Proton-Proton-Reaktionen produziert. Infolge der schlechten Durchmischung des Sterninneren bildet sich im Laufe der Zeit eine „wasserstoffbrennende" Schalenquelle heraus, die einen „ausgebrannten" Heliumkern umgibt. Bevor es allerdings zum „Heliumbrennen" kommen kann, muß die Temperatur im Sternzentrum auf ungefähr 100 Mill. Kr ansteigen. Dies wird durch eine allmähliche Kontraktion des Zentralgebiets erreicht. Während sich der Stern im Inneren immer mehr verdichtet, expandiert er als ganzes, seine Leuchtkraft nimmt dabei immer mehr zu, hauptsächlich infolge der enormen Radienvergrößerung, er wird zu einem *Roten Riesen.*

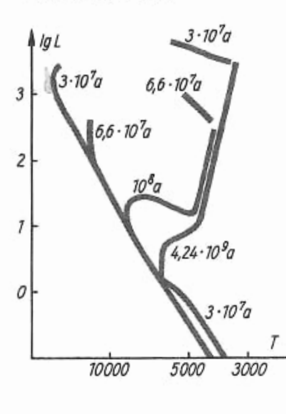
Ähnlich wie im Falle des „Kohlenstoffbrennens" in massereicheren Sternen, setzen die Heliumreaktionen in der hochgradig entarteten Materie schlagartig ein. Wenige Minuten lang entfaltet der Stern in seinem Inneren die Leuchtkraft eines ganzen Sternsystems. Diese extrem kurze Phase wird als *Flash* [*flash* engl., Blitz] bezeichnet. Ein äußerer Beobachter bemerkt davon allerdings nichts, da die so plötzlich erzeugte Energiemenge einige 10^6 Jahre benötigt, um sich in Form von Strahlung durch den Stern nach außen „durchzuarbeiten". Die Entwicklungsrechnungen konnten bisher noch nicht über die Flash-Phase hinaus weitergeführt werden, da sich die physikalische Instabilität im Sternzentrum nur schwer numerisch modellieren läßt. Der weitere Entwicklungsweg ist deshalb hypothetisch.

Im Unterschied zur Kohlenstoffdetonation, die der weiteren Entwicklung massereicher Sterne ein Ende setzt, folgt auf den Helium-Flash der sonnenähnlichen Sterne wieder eine ruhige Entwicklungsphase, in der die „heliumbrennende" Kernquelle von einer den Wasserstoffvorrat der weiter außen liegenden Schichten verwertenden Schalenquelle umgeben ist. Der Stern bewegt sich dann entlang des sogenannten Horizontalasts im Hertzsprung-Russel-Diagramm, wobei eventuell auch der Instabilitätsbereich der RR Lyrae-Sterne, der massearmen Pulsationsveränderlichen der Population II (↑ 10.3.5.1.), gekreuzt wird. Zum „Kohlenstoffbrennen" kommt es bei Sternen unter einer Sonnenmasse nicht, da die Masse der sich bildenden Kohlenstoffkugel einfach nicht ausreicht, um die zum Zünden erforderrlichen Werte für Temperatur und Dichte zu erreichen.

In den späten Entwicklungsphasen der Sterne spielen *Masseverluste* wahrscheinlich eine entscheidende Rolle. So können z. B. in den kühlen Atmosphären von Riesensternen Staubteilchen, wie Silicate oder Graphit, auskondensieren und durch den Strahlungsdruck des Sterns in den interstellaren Raum transportiert werden. Der unregelmäßige Veränderliche R. Coronae-Borealis, der von Zeit zu Zeit plötzliche Helligkeitseinbrüche von mehreren Größenklassen in seiner Lichtkurve zeigt und Prototyp einer nach ihm benannten Klasse von Veränderlichen ist, stellt wahrscheinlich einen solchen „rußenden" Riesenstern dar, in dessen Atmo-

sphäre Wasserstoff nur noch in Spuren vorkommt und Helium und Kohlenstoff die Atmosphärenchemie bestimmen. Diese merkwürdige Atmosphärenzusammensetzung kann mit einer starken Durchmischung des gesamten Sterns erklärt werden, in deren Folge die Abbrandprodukte Helium und Kohlenstoff nach oben gespült wurden und der Wasserstoff der Hülle im heißen Sterninneren thermonuklear verbrannt wurde.

Eine weitere Möglichkeit der Überprüfung der Theorie bieten enge Doppelsternsysteme, bei denen die Entwicklung beider Komponenten durch gegenseitigen Masseaustausch zum Teil erheblich modifiziert wird. Jeder Stern durchläuft während seiner Entwicklung expansive Phasen. Überschreitet dabei sein Volumen den durch die *Rochesche Grenze* gesetzten kritischen Wert, fließt Materie zum Begleiter über. Auf diese Art und Weise kann es zu einem beträchtlichen Masseaustausch kommen, in dessen Verlauf sich sogar das ursprüngliche Masseverhältnis zwischen den beiden Sternen umkehren kann. Zur Illustration sei ein Rechenmodell angeführt, das die Entwicklung eines engen Doppelsternsystems, bestehend aus zwei Hauptreihensternen von 9 und 3,1 Sonnenmassen, die zu Beginn 29,5 Sonnenradien voneinander entfernt sind, rechnerisch nachvollzieht (Abb.



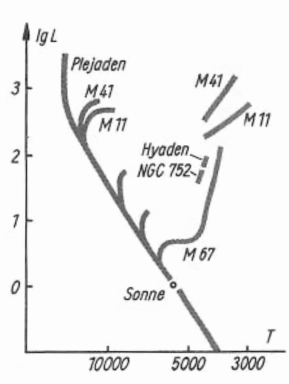


Abb. 10.3.3-9 Vergleich verschiedener Entwicklungsstadien eines theoretischen Sternhaufens (nach KIPPENHAHN und WEIGERT) mit dem Hertzsprung-Russell-Diagramm von beobachteten, unterschiedlich alten Sternhaufen; L Leuchtkraft in Sonneneinheiten, T effektive Temperatur in K, a Jahre

10.3.3.8.3. *Prüfung der Theorie.* Die Entwicklung selbst der massereichsten Sterne ist so langsam, daß keine beobachtbaren Änderungen von Leuchtkraft oder effektiver Temperatur in irdischen Zeiträumen erfolgen. Zur Prüfung der Theorie benutzt man Sternhaufen; da alle Sterne in ihnen gleichzeitig im gleichen Raumgebiet aus interstellarer Materie entstanden sind, hatten sie gleiche chemische Zusammensetzung bei Beginn der Kernprozesse und haben gleiches Alter, aber unterschiedliche Masse. Wenn die theoretischen Entwicklungswege richtig sind, entfernen sich die Sterne verschiedener Masse unterschiedlich schnell auf ihren spezifischen Entwicklungswegen von der Hauptreihe, und für gleichaltrige Sterne mit gleicher ursprünglicher Zusammensetzung lassen sich Linien konstruieren, auf denen sie nach einer bestimmten Zeit liegen müssen. Ähnliche Linien müssen sich ergeben, wenn man die beobachteten Sterne eines Sternhaufens im Hertzsprung-Russell- oder im Farben-Helligkeits-Diagramm einträgt. Nach dem Vergleich verschiedener Entwicklungsstadien eines konstruierten mit beobachteten Sternhaufen (Abb. 10.3.3-9) können bei guter Übereinstimmung Altersangaben für beobachtete Sternhaufen gemacht werden: je jünger der Sternhaufen ist, um so weiter ist die Hauptreihe in Richtung auf frühe Spektralklassen besetzt. Die Plejaden dürften demnach ein Alter von etwa $5 \cdot 10^7$ Jahren haben, der Sternhaufen M 67 hingegen dürfte $5 \cdot 10^9$ Jahre alt sein.

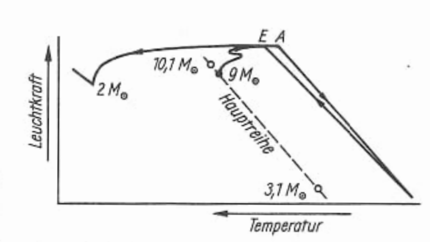


Abb. 10.3.3-10 Enwicklung eines Doppelsternsystems von 9 und 3,1 Sonnenmassen; der 9-Sonnenmassen-Stern ist auf dem Weg zum Roten Riesen, wenn im Punkt A der Massenaustausch zwischen beiden Sternen beginnt; E Ende des Masseverlustes; der ursprünglich masseärmere Stern befindet sich auf der Hauptreihe als 10,1-Sonnenmassen-Stern, der übriggebliebene Heliumstern von 2 Sonnenmassen bewegt sich nach dem Einsetzen des „Heliumbrennens" schnell nach links und verbleibt auf der Heliumreihe (nach KIPPENHAHN und WEIGERT)

10.3.3-10). Der 9-Sonnenmassen-Stern entwickelt sich sehr schnell zum Roten Riesen und erreicht, noch bevor das „Heliumbrennen" einsetzt, seine Rochesche Grenze. Durch den jetzt einsetzenden rapiden Masseaustausch verliert er in nur 40000 Jahren nahezu seine ganze wasserstoffreiche Hülle, d. h. 7 Sonnenmassen. Übrig bleibt ein entwickelter heißer Heliumstern mit einer für seine Masse von 2 Sonnenmassen ungewöhnlich hohen Leuchtkraft und ein normaler Hauptreihenstern von nunmehr 10,1 Sonnenmassen, der aus der ursprünglich masseärmeren Komponente hervorgegangen ist. Dies Beispiel zeigt, daß Sterne ungewöhnlicher chemischer Zusammensetzung dadurch entstehen können, daß durch den Verlust der äußeren Hülle Gebiete freigelegt werden, die bereits durch Kernreaktionsrückstände chemisch kontaminiert sind.

10.3.3.9. Endstadien der Sternentwicklung. Die Entwicklung der Sterne vollzieht sich in Etappen. Ruhige Entwicklungsphasen, die durch einen bestimmten Typ der Energieerzeugung gekennzeichnet sind, wechseln mit kurzen Phasen ab, in denen das Sterninnere kontrahiert und bei ansteigender Zentraltemperatur neue Energiequellen erschlossen werden.

Sind schließlich alle verwertbaren Energievorräte erschöpft, das kann bei massereichen Sternen bereits nach einigen 10^7 Jahren der Fall sein, und ist der Stern nicht der Kohlenstoffdetonation zum Opfer gefallen, schrumpft er erneut unter seinem Eigengewicht zusammen. Sein weiteres Schicksal hängt nunmehr entscheidend von seiner, eventuell durch Masseverluste erheblich dezimierten Restmasse ab.

10.3.3.9.1. *Weiße Zwerge.* Ist die Masse der kollabierenden Sterne kleiner als die von CHANDRASEKHAR angegebene Grenzmasse von ungefähr 1,4 Sonnenmassen, wird die gravitative Kontraktion bei Zentraldichten von 10^5 bis 10^8 gcm^{-3} vom Druck des entarteten Elektronengases aufgefangen (↑ 8.4.3.4.). Der Stern hat dann planetare Ausmaße erreicht und wird als Weißer Zwerg bezeichnet. Bei vorgegebener chemischer Zusammensetzung besteht für solche selbstgravitierenden entarteten Gaskugeln eine eindeutige Beziehung zwischen Masse und Sternradius (Abb. 10.3.3-11). Anders als bei normalen Sternen verringert sich der Radius mit zunehmender Masse, um bei dem kritischen Massewert gegen Null zu streben. Die Temperatur spielt für den inneren Aufbau der Weißen Zwerge keine Rolle, da bei hochgradig entarteter Materie der Druck nur noch von der Elektronendichte abhängt.

Im Hertzsprung-Russell-Diagramm bevölkern die Weißen Zwerge den leuchtschwachen, unteren Bereich (↑ Abb. 10.3.1-7). Infolge ihrer kleinen strahlenden Oberfläche ist ihre absolute Helligkeit trotz hoher effektiver Temperaturen nur gering, sie sind deshalb auch nur in der näheren Sonnenumgebung zu beobachten. Heute sind etwa 1000 Weiße Zwerge bekannt, viele davon als lichtschwache Komponenten von Doppelsternen. Der erste Weiße Zwergstern wurde bereits 1852 im Siriussystem (Sirius B) entdeckt, aber erst 1914 als solcher erkannt. (Über die mögliche Entstehung von Weißen Zwergen in engen Doppelsternsystemen s. a. 10.3.3.8.)

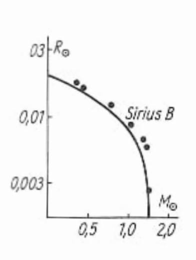


Abb. 10.3.3-11 Theoretische Masse-Radius-Beziehung für entartete Gaskugeln; die Punkte entsprechen beobachteten Weißen Zwergen; Massen in Sonnenmassen, Radien in Sonnenradien (nach SCHATZMAN)

Da die Ausstrahlung der Weißen Zwerge nur durch den Vorrat an innerer Wärme gedeckt wird, kühlen sie langsam, bei nahezu konstantem Radius, aus. Nach vielen Jahrmilliarden wird so aus einem Weißen zunächst ein *Roter* und schließlich ein *Schwarzer Zwerg*, der sich der Beobachtung völlig entzieht. Schwarze Zwergsterne entstehen wahrscheinlich auch dann, wenn die Masse eines Protosterns nicht zum Zünden des „Wasserstoffbrennens" ausreicht. Für Hauptreihensterne besteht somit eine natürliche untere Massegrenze, die, etwas von der chemischen Zusammensetzung abhängig, bei etwa 0,07 Sonnenmassen oder 70 Jupitermassen liegen dürfte.

10.3.3.9.2. *Neutronensterne.* Überschreitet die Masse des kollabierenden Sterns die Chandrasekhargrenze, vermag der Druck der freien Elektronen den Gravitationskollaps nicht aufzuhalten, und der Stern erreicht erst bei einer Dichte von 10^{15} gcm^{-3} (1 Mrd. tcm^{-3}), wenn der Entartungsdruck der Neutronen bedeutsam wird, einen stabilen Endzustand. Bei einer derartigen Packungsdichte der Materie – sämtliche Hüllenelektronen sind längst in die Atomkerne „hineingequetscht" worden – ähnelt der gesamte Stern einem riesigen, neutronenangereicherten Atomkern. Solche Neutronensterne sind nur etwa 10 bis 15 km groß und rotieren mit großer Geschwindigkeit um ihre Rotationsachse, da der ursprüngliche Drehimpuls des Sterns während seines Kollapses erhalten bleibt, man spricht vom Pirouetteneffekt.

Gleich nach ihrer Entdeckung durch britische Radioastronomen 1967/68 wurden die *Pulsare* mit den bislang erfolglos gesuchten Neutronensternen, auf deren mögliche Existenz schon

30 Jahre zuvor unter anderem von LANDAU hingewiesen wurde, in Verbindung gebracht. Heute sind bereits 160 dieser merkwürdigen Radioquellen bekannt, die mit konstantem, für jeden Pulsar charakteristischen Zeitabstand Impulse im Radiofrequenzbereich aussenden. Die gemessene Pulsfrequenz kann bis zu 30 Impulsen pro Sekunde betragen. Langperiodische Pulsare, die nur alle paar Sekunden einen Strahlungsausbruch zeigen, sind wegen ihrer mit zunehmender Periodenlänge immer geringer werdenden Radioleuchtkraft nur schwer aufzufinden. Bemerkenswerterweise beschränkt sich die Variabilität bei zwei jungen Pulsaren, dem Krebsnebelpulsar NP 0532 und dem Velapulsar PSR 0833-45, nicht nur auf die Radiostrahlung, sondern ist sowohl im optischen als auch im Röntgenbereich mit derselben Periode nachweisbar.
Die außergewöhnliche Konstanz der gemessenen Periodenlängen, in vielen Fällen sind die Perioden bis auf 10^{-9} s genau bekannt, kann nur durch die schnelle Rotation eines kompakten Objekts, eines Neutronensterns, erklärt werden.
Die Strahlung ist nichtthermischer Natur (Synchrotronstrahlung) und entsteht wahrscheinlich in dem den Neutronenstern umgebenden Plasma in der Nähe des sogenannten Lichtzylinders. Hier, im Falle des Krebsnebelpulsars nur 1 500 km von der Pulsaroberfläche entfernt, erreicht nämlich sowohl das mit dem Pulsar mitrotierende Magnetfeld als auch das vom Magnetfeld eingefangene Plasma fast Lichtgeschwindigkeit und infolgedessen sind die Bedingungen für das Aussenden von *Synchrotronstrahlung* und von *relativistischen Teilchen* besonders günstig. Warum die Strahlung eng gebündelt wie ein Scheinwerferstrahl den Himmel überstreicht und bei einem entfernt stehenden Beobachter den Eindruck einer periodisch flakkernden Lichtquelle erweckt, man spricht auch vom *Leuchtturmeffekt*, ist in allen Einzelheiten noch nicht verstanden. Gewisse Vorzugsrichtungen könnten dadurch zustande kommen, daß die Symmetrieachse des fest mit dem Neutronenstern verankerten Magnetfelds gegen die Rotationsachse des Sterns geneigt ist. Magnetfelder spielen mit Sicherheit eine große Rolle, da bereits schwache stellare Magnetfelder, wie sie von der Sonne und anderen Sternen bekannt sind, während des Sternkollapses in der elektrisch leitenden Sternmaterie „eingefroren" sind und schon deshalb auf enorme Flußdichten von größenordnungsmäßig 10^8 T verstärkt werden müssen.
Zwei Pulsare wurden bisher in unmittelbarer Nachbarschaft von *Supernovaüberresten* gefunden und legen die Deutung nahe, in ihnen die extrem verdichteten Überbleibsel von explodierenden massereichen Sternen zu sehen. Für einen solchen genetischen Zusammenhang spricht im Falle des Krebsnebelpulsars auch die Tatsache, daß die für das Nebelleuchten und die beschleunigte Expansion des Nebelgases benötigte Energie offensichtlich vom Pulsar nachgeliefert wird. Anders ist die gemessene langsame Zunahme der Periodenlänge von $35 \cdot 10^{-9}$ s pro Tag, die einem Verlust an Rotationsenergie von ungefähr dem geforderten Betrag von $4 \cdot 10^{31}$ W entspricht, nicht zu erklären. Bereits nach 2500 Jahren wird sich die Rotationsdauer dieses Pulsars als Folge der Dissipation seiner Rotationsenergie verdoppelt haben. Auch von anderen Pulsaren sind geringfügige Periodenverlängerungen bekannt, manchmal treten allerdings auch sprunghafte Periodenverkürzungen auf, die mit „tektonischen" Vorgängen innerhalb der starren Kruste eines Neutronensterns, sogenannten Sternbeben, erklärt werden können.
Die Pulsare sind für den Astrophysiker aus verschiedenen Gründen von Interesse. Einmal gestattet ihr Studium Einblick in das Verhalten der Materie unter extrem hohen Drücken, die unter Laborbedingungen niemals erreichbar sein werden, andererseits kommt ihnen als Quelle der kosmischen Strahlung große Bedeutung für den Energiehaushalt der Galaxis zu, und schließlich dienen sie als Testobjekte für die verschiedenen Gravitationstheorien, deren theoretische Vorhersagen sich erst bei den enorm hohen Gravitationsfeldstärken in der Umgebung von Neutronensternen nennenswert voneinander unterscheiden.
Auch die Erforschung der interstellaren Materie profitiert von den Pulsarbeobachtungen. Aus den gemessenen Frequenzdriften in den Ankunftszeiten der Radioimpulse – ein und derselbe Pulsarimpuls kann infolge der Dispersion der Radiowellen im interstellaren Medium im langwelligen Radiobereich mit bis zu mehreren Minuten Verspätung gegenüber den im kurzwelligen Radiobereich beobachteten Impuls eintreffen (↑ 2.3.1.4.) – und der Drehung der Polarisationsebene der Pulsarstrahlung in Abhängigkeit von der Beobachtungsfrequenz, dem *Faradayeffekt*, können mittlere Elektronendichten und Magnetfeldstärken im interstellaren Raum bestimmt werden (↑ 9.1.5.2.3.).
Neutronensterne als Komponenten enger Doppelsternsysteme werden auch herangezogen, um die von Satelliten entdeckten *pulsierenden Röntgenquellen* zu erklären. Beispielsweise zeigt der „Röntgenstern" Hercules X-1, dessen Röntgenhelligkeit mit einer Periode von 1,24 s flackert, einen typischen Bedeckungslichtwechsel. Alle 1,7 Tage wird für mehrere Stunden die Röntgenquelle, die die masseärmere Komponente eines engen Doppelsternsystems ist, durch den Hauptstern bedeckt. Die Quelle ist mit dem schon lange bekannten Veränderlichen HZ Herculis identisch, dessen optische Helligkeit mit derselben Periode variiert. Die beiden Sterne sind

$4 \cdot 10^6$ km voneinander entfernt, die Gesamtmasse des Systems beläuft sich auf 3,4 Sonnenmassen. Die Bahndaten sind aus Dopplereffektmessungen sehr genau bekannt, weil sich durch die Bewegung der Röntgenquelle um den Systemschwerpunkt die Ankunftszeiten der Röntgenpulse während eines Umlaufs systematisch ändern – bei Annäherung wird eine kürzere Periode registriert als wenn sich die Quelle von uns entfernt – und sich aus der gemessenen Lichtlaufzeit die absoluten Dimensionen des Systems leicht berechnen lassen.

Die Röntgenstrahlung ist thermischen Ursprungs und entsteht wahrscheinlich dadurch, daß der Hauptstern Materie verliert, die von dem begleitenden Neutronenstern eingefangen wird und mit hoher Geschwindigkeit auf den Neutronenstern „abregnet". Beim Aufprall wird das Gas auf einige 10^7 K erhitzt. Das Maximum der entweichenden thermischen Strahlung liegt bei einem derart hoch erhitzten Plasma im Röntgenbereich.

Unter der Annahme, daß die überströmende Materie längs der magnetischen Feldlinien des schnell rotierenden Neutronensterns bevorzugt zu den Magnetpolen geführt wird, lassen sich auch die beobachteten Röntgenpulse leicht verstehen: die durch den kontinuierlichen Gaseinfall stark aufgeheizten Polregionen geraten als sogenannte heiße Flecke periodisch in das Blickfeld des Beobachters.

Der optische Lichtwechsel des Veränderlichen findet seine zwanglose Erklärung in der zusätzlichen Aufheizung der der Röntgenquelle zugewandten Seite des Hauptsterns.

Über den Strahlungsmechanismus einer erst kürzlich entdeckten Gruppe von galaktischen Röntgenquellen, die mit einer gewissen Regelmäßigkeit, meist im Abstand einiger Stunden, kurzzeitige Strahlungsausbrüche im Röntgen- und γ-Bereich zeigen, gibt es bisher nur Hypothesen. Einige dieser sogenannten *Bursters* [*burst* engl., bersten] wurden in Kugelsternhaufen gefunden. Während eines nur wenige Sekunden bis Minuten dauernden Ausbruchs wird allein im Röntgenbereich eine Energiemenge freigesetzt, die der Gesamtstrahlung der Sonne im Laufe eines ganzen Jahres entspricht. Der schnellste bisher beobachtete Burster, MXB 1730-335, „feuert" etwa 5000mal am Tag, wobei besonders starke Ausbrüche immer von einer relativ langen Zeit der Ruhe gefolgt werden (Abb. 10.3.3-12).

Ein gängiger Erklärungsversuch geht davon aus, daß die Magnetosphäre eines Neutronensterns eine bestimmte Menge heißen Gases speichern kann und solange zurückbehält, bis sich die angesammelte Gasmasse entweder hinreichend abgekühlt hat oder einfach durch ihr zunehmendes Eigengewicht plötzlich durch die Magnetosphäre hindurch auf die Oberfläche des Neutronensterns stürzt und dabei die während des freien Falls gewonnene Energie in Strahlung umsetzt. Dieser Strahlungsausbruch heizt nun wiederum die einfallende Materie so stark auf, daß sie wieder für eine Zeitlang von der Magnetosphäre zurückgehalten werden kann, bis es zu einem neuen Ausbruch kommt. Das einfallende Gas kann wie bei den Röntgendoppelsternen von einem nahen Begleiter stammen.

10.3.3.9.3. *Schwarze Löcher.* Während Neutronensterne als Pulsare und Röntgensterne der direkten Beobachtung zugänglich sind, kann dies von einer noch „exotischeren" Klasse kondensierter Objekte, den Schwarzen Löchern, noch nicht mit Bestimmtheit gesagt werden. Die theoretischen Überlegungen, die die Existenz der Schwarzen Löcher wahrscheinlich werden lassen, gehen davon aus, daß auch Neutronensterne nur bis zu einer Masse von zwei Sonnenmassen stabil sein können. Selbst die abstoßenden Kernkräfte sind nicht in der Lage, den Gravitationskollaps massereicherer Sterne aufzuhalten. Sie stürzen unaufhaltsam in sich zusammen, wobei ihr Gravitationsfeld schließlich so stark wird, daß sich ein Stern bei der Annäherung an seinen *Schwarzschildradius* $r_s = 2GM/c^2$ (G Gravitationskonstante, M Masse des Sterns, c Lichtgeschwindigkeit) von der übrigen Welt abschnürt, da kein noch so energiereiches Teilchen ihn mehr verlassen kann. Obwohl der Schwarzschildradius eine charakteristische Größe der Einsteinschen Gravitationstheorie ist, kann er anschaulich als derjenige Radius gedeutet werden, den ein Himmelskörper vorgegebener Masse haben müßte, damit die Entweichgeschwindigkeit von seiner Oberfläche gleich der universellen Lichtgeschwindigkeit

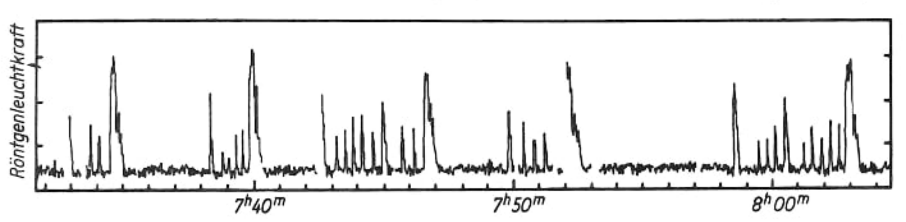


Abb. 10.3.3-12 Röntgenstrahlungsausbrüche des schnellen „Bursters" MXB 1730 – 335 im Verlaufe einer halben Stunde (nach LEWIN)

wird. Setzen wir die Masse der Sonne, d. h. $2 \cdot 10^{33}$ g, in die obige Formel ein, errechnet sich der Schwarzschildradius der Sonne zu nahezu 3 km, das ist ein fünfmillionstel des tatsächlichen Sonnenradius. Im Falle der Erde ergibt sich sogar nur ein Durchmesser von 11 mm.

Bei den bereits erwähnten Durchmusterungen des Himmels nach kompakten Röntgenquellen stieß man auch auf Doppelsterne, bei denen die Masse der Röntgenkomponente die theoretische Massengrenze für stabile Neutronensterne überschreitet. Für den kompakten Begleiter des Röntgendoppelsterns *Cygnus X-1* ergibt sich z. B. eine Masse von über 6 Sonnenmassen. Falls dieser Massewert richtig ist und unsere theoretischen Vorstellungen über die Stabilität von Neutronensternen korrekt sind, kann es sich bei diesem Objekt nur um ein Schwarzes Loch handeln. Man wird abwarten müssen, inwiefern weitere Beobachtungen diese Schlußfolgerung bestätigen.

Die Stellung der Schwarzen Löcher [engl. *black holes*] im Rahmen der modernen Physik ist umstritten. Der mit der Existenz der Schwarzen Löcher verbundene prinzipielle Informationsverlust über alle physikalischen Prozesse, die sich innerhalb des Ereignishorizonts $r < r_S$ abspielen, erlaubt einens pekulativen Brückenschlag zur Thermodynamik, da der Schwarzschildradius, ähnlich wie die Entropie eines thermodynamischen Systems, nur gleichbleiben oder durch Massezufluß anwachsen, sich aber niemals verringern kann. Fernerhin verknüpft die mikroskopische Kleinheit selbst schwerer Schwarzer Löcher von Asteroidenausmaß die Gravitations- mit der Quantentheorie. Beispielsweise sollten sogenannte „Minilöcher" [engl. *mini holes*] die evenutell in der Frühphase des Kosmos in großer Zahl entstanden sein können, über den quantenmechanischen Tunneleffekt sogar Teilchen emittieren, also nicht völlig „schwarz" sein.

Diese vielfältigen Anknüpfungspunkte der Theorie der Schwarzen Löcher mit sonst so wesensverschiedenen Gebieten der Physik, wie der Gravitationstheorie, der Quantenmechanik und der phänomenologischen Thermodynamik, erwecken gewisse Hoffnungen, einer einheitlichen Physik näherzukommen.

10.3.4. Interstellare Materie

Im Milchstraßensystem ist der Raum auch zwischen den Sternen mit Materie angefüllt, deren Gesamtmasse zwar nur etwa 5% der in den Sternen konzentrierten Masse ausmacht, der aber dennoch große kosmogonische Bedeutung zukommt. Auch heute bilden sich noch aus diesem extrem verdünnten Gas – die mittlere Materiedichte im interstellaren Raum ist mit 10^{-24} gcm^{-3} viel geringer als in jedem technisch erzeugten Ultrahochvakuum – Sterne, die in späten Phasen ihrer Entwicklung auch wieder Materie, inzwischen durch die stellaren Energieerzeugungsprozesse mit schweren Elementen angereichert, in den interstellaren Raum abgeben, so daß ein ständiger Austausch zwischen der stellaren und der interstellaren Materie stattfindet.

Die *interstellare Materie*, bestehend aus Gas, dem in geringer Menge feste *Staubteilchen* beigemischt sind, kann nicht losgelöst vom *interstellaren Strahlungsfeld, magnetischen Feldern* und der *kosmischen Teilchenstrahlung* betrachtet werden, da der physikalische Zustand und die Bewegungsverhältnisse durch die Wechselwirkung mit diesen Feldern bestimmt sind.

Die interstellare Materie ist in wolkigen Verdichtungen konzentriert und kann leuchtend als *Emissions-* oder *Reflexionsnebel*, nichtleuchtend als *Dunkelwolke* oder nur als im Radiofrequenzbereich beobachtbare *HI-Wolke* auftreten. Als typischer Vertreter der extremen Population I (↑ 10.3.5.1.) ist die interstellare Materie sehr stark zur galaktischen Ebene konzentriert und bildet eine extrem flache Scheibe mit einer Dicke von nur etwa 220 pc (1 pc = $30{,}857 \cdot 10^{15}$ m). In der Scheibe wiederum sind die interstellaren Gas- und Staubwolken bevorzugt längs der Spiralarme unserer Galaxis angeordnet. Die Wolken haben typische Durchmesser von 5 bis 40 pc, die Masse einer Wolke kann bis zu $2 \cdot 10^4$ Sonnenmassen betragen, doch sind die massereichen Wolken relativ selten. In den Wolken liegt die mittlere Teilchendichte bei ungefähr 10 Wasserstoffatomen pro cm^3, es kommen allerdings auch beträchtlich höhere Werte vor. Die Wolken scheinen sich im Druckgleichgewicht mit einem heißen, aber hochverdünnten Zwischenwolkengas mit einer Teilchendichte von nur 0,05 pro cm^3 zu befinden. Die hohe Temperatur dieses Gases von schätzungsweise 10^4 K ist eine Folge der Aufheizung durch die weiche Komponente der kosmischen und der Röntgenstrahlung.

10.3.4.1. Interstellares Gas. Die Atome, Ionen, Elektronen und Moleküle im interstellaren Raum bilden das interstellare Gas; es besteht hauptsächlich aus Wasserstoff, dem im Verhältnis 10 : 1 Helium beigemischt ist. Die anderen Elemente sind außerordentlich selten. Der interstellare Wasserstoff ist in den *HI-Gebieten* neutral, in den *HII-Gebieten* aber durch die ultraviolette Sternstrahlung mit Wellenlängen kleiner als 91,2 nm, was der Lymangrenze entspricht, ionisiert. Diese Strahlung stammt von den seltenen Sternen der frühen Spektralklassen mit hoher effektiver Temperatur und hoher Leuchtkraft, die sich in unmittelbarer Nähe des leuchtenden Gasnebels aufhalten.

Die HI-Gebiete treten optisch nur durch die dunklen Absorptionslinien, die den Spektren entfernter Sterne zusätzlich durch das interstellare Gas aufgeprägt werden, in Erscheinung. Die stärksten interstellaren Absorptionslinien stammen vom einfach ionisierten Calcium und neutralen Natrium. Jedoch sind auch Linien anderer zum Teil einfach ionisierter Atome und sogar Moleküllinien gefunden worden. Mit nur einer Ausnahme, den Linien des Cyanmoleküls, entstehen alle interstellaren Linien durch Übergänge vom Grundzustand aus.
Die Energiedichte des interstellaren Strahlungsfelds ist so gering, daß trotz hoher Farbtemperatur, im ultravioletten Bereich bis zu 30000 K, die Absorptionen so selten erfolgen, daß sich alle Atome und Moleküle in ihrem energetischen Grundzustand aufhalten.
Die Anregung des Cyanmoleküls fand erst spät seine Erklärung durch die Entdeckung der universellen *3-Kelvin-Strahlung* (↑ 10.4.4.1.), mit der sich das Molekül im thermodynamischen Gleichgewicht befindet. Nicht alle im interstellaren Gas vorhandenen Elemente werden beobachtet; gerade die häufigsten Elemente wie H, He, O, C und N treten nicht in Absorption auf. Erst mittels Satelliten ist es möglich geworden, auch die Linien dieser Elemente, die im ultravioletten Spektralbereich liegen und daher von der Erde aus nicht beobachtet werden können, nachzuweisen und auch die chemischen Häufigkeiten dieser Elemente zu bestimmen. Dabei zeigt sich, daß im HI-Gas die Metalle relativ zu ihren Häufigkeiten in den Sternatmosphären der Population-I-Sterne unterhäufig sind. Dieses Metalldefizit kann zum Teil dadurch erklärt werden, daß die schweren Elemente in Form von kosmischen Staubteilchen auskondensieren und damit dem Gas entzogen werden.
Die interstellaren Absorptionslinien sind vielfach aufgespalten und die Komponenten sind gegeneinander verschoben. Aufgrund des Dopplereffekts bedeutet das unterschiedliche Radialgeschwindigkeiten und damit großräumige Bewegungen der HI-Wolken.
Einen großen Fortschritt in der Beobachtung der HI-Gebiete brachte die Radioastronomie, da der neutrale interstellare Wasserstoff im Radiofrequenzbereich im Unterschied zum optischen Spektralbereich direkt beobachtbar ist. Der Grundzustand des Wasserstoffs hat eine Hyperfeinstruktur (↑ 4.2.4.8.). Elektronenspin und Kernspin können parallel oder antiparallel sein. Beim Übergang von einem zum anderen Zustand wird Strahlung der Wellenlänge 21 cm ausgesandt. Infolge der radialen Wolkenbewegungen ist die beobachtete *21-cm-Linienemission* in mehrere Komponenten aufgespalten (Abb. 10.3.4-1). Die Unterschiede in den gemessenen Radialgeschwindigkeiten der einzelnen Wolken sind im wesentlichen eine Folge der differentiellen Rotation des Milchstraßensystems. Bei bekanntem Rotationsgesetz ist es möglich, aus den Radialgeschwindigkeiten auf die Entfernungen der Wasserstoffwolken zu schließen und so ein Bild der Verteilung des neutralen Wasserstoffgases in der Milchstraßenebene zu gewinnen (↑ 10.3.5.3.; Tafel 62).
Die geringe kinetische Temperatur der HI-Gebiete von unter 200 K ist eine Folge der effektiven Kühlung des Gases durch Stoßanregung niederenergetischer Niveaus des Sauerstoffs, Kohlenstoffs, Siliciums und anderer schwerer Elemente. Die bei der Rückkehr auf das Grundniveau emittierte Strahlung verläßt die Wolke, die für diese im Infrarotbereich liegende Strahlung optisch dünn ist. Auch die Anregung von Rotationsniveaus des H_2-Moleküls und die thermische Strahlung der Staubteilchen führt zu Energieverlusten und damit zu einer Kühlung

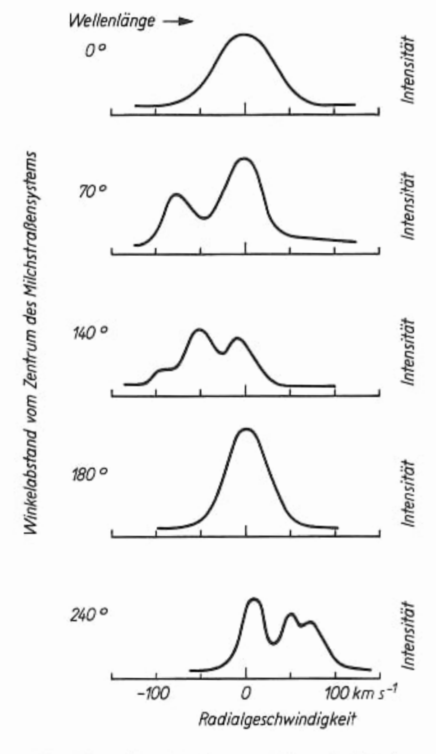


Abb. 10.3.4-1 Profile der 21-cm-Linie in verschiedenen Richtungen innerhalb der Milchstraße; die Intensität ist statt über der Wellenlänge über der Dopplergeschwindigkeit aufgetragen; einer Radialgeschwindigkeit von 100 km s^{-1} entspricht eine Wellenlängenänderung von 0,07 mm

des Gases. Diesen Kühlprozessen steht die gelegentliche sehr starke Aufheizung der Wolken durch Wolkenstöße oder Supernovaausbrüche entgegen. Die kinetische Energie der stoßenden Wolken bzw. der expandierenden Supernovahülle wird dabei über Stoßwellen nahezu vollständig thermalisiert, wobei die auftretenden Geschwindigkeiten die Schallgeschwindigkeit beträchtlich überschreiten.

Nur wenige Prozent des interstellaren Gases befinden sich im Einflußbereich der Strahlung heißer Sterne der frühen Spektralklassen und können als *leuchtende Gasnebel* beobachtet werden. Die Ausdehnung dieser HII-Gebiete ist um so größer, je höher die effektive Temperatur der leuchtanregenden Sterne ist. So kann ein 05-Stern ein Gebiet von etwa 300 pc Durchmesser durch seine UV-Strahlung ionisieren. ein A0-Stern hingegen nur noch eine Wasserstoffsphäre von 3 pc Ausdehnung. Ein ausgedehntes und in seinem Inneren sehr dichtes HII-Gebiet stellt der Orionnebel dar (Tafeln 64 und 65). Da in den HII-Gebieten praktisch aller Wasserstoff ionisiert ist, ist die Zahl der freien Elektronen gleich der der Protonen. In den HII-Gebieten stellt sich ein Gleichgewicht ein zwischen der Zahl der Ionisationen mit Energieaufnahme aus dem Strahlungsfeld der umgebenden Sterne und der Anzahl der Rekombinationen mit Energieabgabe. Im Gleichgewicht herrscht eine mittlere kinetische Temperatur von rund 10000 K. Bei Energieabgabe werden Spektrallinien emittiert, die das HII-Gebiet sichtbar machen. Bei Dichten bis zu 1000 Teilchen je cm^3 erscheint es als hell leuchtender *Gasnebel*. Die stärksten Emissionslinien stammen von H, He, He^+, O^+, O^{++}, N^+ und Ne^{++}. Das *Nebelleuchten* ist nicht thermischen Ursprungs wie bei den Sternen; die ermittierte Energie stammt vielmehr aus dem Licht benachbarter heißer Sterne. Neben der Rekombination wirkt die Stoßanregung von tiefliegenden Energieniveaus einiger Ionen und Atome durch Elektronen. Dabei werden die *Nebellinien*, verbotene Linien von O^+, O^{++} und N^+, angeregt. Die Anregung erfolgt auf metastabile Niveaus, deren Übergangswahrscheinlichkeiten viel kleiner sind als die von normalen Niveaus. Im Laboratorium kann es nicht zur Ausstrahlung dieser Linien kommen, da dort weitere Stoß- oder Strahlungsanregungen auf höhere, normale Niveaus erfolgen; im interstellaren Gas erfolgt erst nach Wochen ein Zusammenstoß mit einem anderen Teilchen; Strahlungsabsorption ist noch seltener. Im Radiofrequenzbereich konnten Linien beobachtet werden, die z. B. Elektronenübergängen des Wasserstoffatoms vom 167. auf das 166. und vom 105. auf das 104. Niveau entsprechen. Neben dieser Linienstrahlung hochangeregter Wasserstoffatome wird radioastronomisch auch eine Kontinuumsstrahlung empfangen, die dadurch entsteht, daß freie Elektronen im Coulombfeld der Protonen nicht eingefangen, sondern nur abgelenkt werden und dabei elektromagnetische Wellen emittieren. Diese *thermische Radioemission* tritt besonders im Zentimeterbereich auf. Das gleichfalls vorhandene längerwellige Radiokontinuum ist nichtthermischen Ursprungs und entsteht bei der Wechselwirkung der geladenen Partikeln der kosmischen Strahlung mit dem interstellaren Magnetfeld.

Im interstellaren Gas können eine Reihe *interstellarer Moleküle* beobachtet werden, die sich durch ihre charakteristischen Absorptions- bzw. Emissionslinien bemerkbar machen. Die meisten der heute bekannten rund 40 Moleküle wurden im Radiofrequenzbereich gefunden. Die Linien entsprechen häufig Übergängen zwischen verschiedenen Rotationszuständen der betreffenden Moleküle. Die wichtigsten im interstellaren Gas gefundenen Moleküle sind: H_2, HD, CO, CN, CH, OH, NH_3, SiO, H_2O, HCN. Daneben werden aber auch zum Teil schon recht kompliziert aufgebaute organische Verbindungen wie Ameisensäure (HCOOH), Formaldehyd (H_2CO) und Ethylalkohol (CH_3CH_2OH) in kühlen HI-Wolken beobachtet.

Die Tatsache, daß bereits unter den Bedingungen des interstellaren Raums chemische Reaktionen ablaufen können, die eventuell sogar zur Bildung einfacher Aminosäuren führen, ist von großer Bedeutung für die präbiologische Chemie.

Die bei einigen interstellaren Molekülen, beispielsweise dem Hydroxid (OH), gefundenen anomalen Intensitätsverhältnisse zwischen einzelnen Linien deuten auf große Abweichungen vom thermodynamischen Gleichgewicht hin. Vermutlich beruht die Linienstrahlung auf einem Masereffekt und entsteht durch stimulierte Emission (↑ 4.1.3.1.). Einige OH-Quellen sind derart kompakt – die interferometrisch bestimmten Durchmesser belaufen sich auf wenige astronomische Einheiten –, daß sie mit *protostellaren Wolken*, die wir im Kollapsstadium beobachten, in Verbindung gebracht werden.

Die interstellaren Molekülinien erlauben auch Rückschlüsse auf die *Isotopenhäufigkeiten* im interstellaren Gas (↑ 10.3.1.8.). Eine spezielle Gruppe der leuchtenden Gasnebel stellen die *planetarischen Nebel* dar, die meist eine relativ regelmäßige Form haben und im Fernrohr oft wie Planetenscheibchen aussehen. Ein bekannter planetarischer Nebel ist der Ringnebel im Sternbild Lyra. Die scheinbaren Nebeldurchmesser betragen bis zu 15′, die wahren Durchmesser rund 0,01 bis 5 pc. Die Nebel werden von *expandierenden Gashülllen* gebildet, deren Massen meist kleiner als 0,2 Sonnenmassen sind. Die spektroskopisch ermittelten

Ausdehnungsgeschwindigkeiten liegen im allgemeinen zwischen rund 10 bis 50 kms^{-1}. Das Nebelgas wird durch einen sehr heißen *Zentralstern* zum Leuchten angeregt. Im Spektrum treten wie bei den normalen Gasnebeln besonders intensiv verbotene Linien von O^+, O^{++} und N^+ in Emission auf. Infrarotbeobachtungen deuten darauf hin, daß die Gashüllen auch einen beträchtlichen Staubanteil enthalten.

Die Zentralsterne der planetarischen Nebel gehören mit effektiven Temperaturen bis zu 10^5 K zu den heißesten Sternen. Möglicherweise handelt es sich um Riesensterne, die sich im letzten Stadium ihrer Entwicklung ihrer dünnen Atmosphärenhülle entledigen, bevor der heiße Kern des Sterns zum *Weißen Zwerg* wird. Die Anzahl der planetarischen Nebel in unserer Galaxis wird auf 20000 bis 50000 geschätzt. Nur 1000 sind aber davon katalogisiert.

10.3.4.2. Interstellarer Staub. Der interstellare Staub ist wie das Gas vorwiegend in Wolken konzentriert. Gas und Staub kommen meist zusammen vor. Obwohl die mittlere Staubdichte nur etwa 10^{-26} gcm^{-3} beträgt, d. h., ein Würfel der Kantenlänge 100 m enthält im Durchschnitt ein mikroskopisch kleines Staubteilchen, macht sich der Staub wegen der Größe der Wolken durch eine Schwächung und spektrale Veränderung des Lichts der hinter einer Gas-Staub-Wolke stehenden Sterne bemerkbar. Befinden sich in der Nähe einer Staubwolke Sterne, so wird deren Licht außerdem noch an den Staubteilchen reflektiert. Die Staubwolke erscheint dann als heller, leuchtender *Reflexionsnebel.* Die durch den Staub bedingte *interstellare Extinktion* kann in der galaktischen Ebene bis zu 8 Größenklassen pro kpc betragen. Starke Extinktion täuscht sternleere Gebiete am Himmel vor, da im Vergleich zu benachbarten Sternfeldern Sterne bestimmter scheinbarer Helligkeiten fehlen. Mit Hilfe von vergleichenden Sternzählungen lassen sich Größen und Dichten der Staubwolken bestimmen und die großräumige Verteilung des interstellaren Staubs in der Sonnenumgebung feststellen. In den *Dunkelwolken* ist die Staubdichte 10- bis 20mal größer als der Mittelwert.

Da die interstellare Extinktion hauptsächlich auf der Streuung des Sternlichts an Teilchen von der Größe der Lichtwellenlänge beruht und kurzwelliges Licht stärker gestreut wird als langwelliges, kommt es beim Durchgang des Lichts durch eine Staubwolke zu einer Verrötung der Sternstrahlung, die Extinktion wirkt selektiv.

Neben einer wellenlängenabhängigen Schwächung wird das Sternlicht beim Durchgang durch eine Staubwolke auch teilweise polarisiert, weil die Staubteilchen länglich und in weiten Gebieten durch die Wirkung interstellarer Magnetfelder von 10^{-10} T und mehr parallel zueinander ausgerichtet sind. Trotz starker lokaler Abweichungen liegen die Polarisationsrichtungen vorwiegend parallel zur Milchstraßenebene, so daß man schließen kann, daß die interstellaren Magnetfeldlinien ebenfalls weitgehend parallel zur Milchstraßenebene verlaufen.

Da das interstellare Verfärbungsgesetz zumindest im optischen Bereich keinerlei Aussagen über die chemische Zusammensetzung der Staubteilchen zuläßt, können darüber bisher keine zuverlässigen Angaben gemacht werden.

Infrarotbeobachtungen von den zirkumstellaren Staubhüllen kühler Überriesen deuten darauf hin, daß die Teilchen aus Silicaten aufgebaut sind. Andere Beobachtungsbefunde sind mit Graphitteilchen vereinbar, die in den Atmosphären kühler Riesensterne auskondensieren können. An diese „feuerfesten" Kondensationskerne lagern sich möglicherweise aus dem interstellaren Gas Atome an, die einen lockeren Mantel von etwa 10^{-5} cm Dicke bilden. Neuere Vorstellungen gehen davon aus, daß der Staub als Nebenprodukt bei der Sternentstehung anfällt und im Stadium der Aufklarung der den Protostern umgebenden dichten Hülle durch einen intensiven Sternwind, wie er von den T-Tauri-Sternen bekannt ist, in den interstellaren Raum abgeblasen wird.

10.3.4.3. Felder. Das *interstellare Strahlungsfeld* setzt sich aus den von den Sternen stammenden Lichtquanten zusammen. Unter mittleren interstellaren Bedingungen würde die hochgradig „verdünnte" Sternstrahlung gerade genügen, einen schwarzen Körper auf einer Temperatur von ungefähr 3 K zu halten. Soll das Strahlungsfeld jedoch nach seiner spektralen Zusammensetzung charakterisiert werden, so ergeben sich für die *Farbtemperatur* der Strahlung je nach dem Wellenlängenbereich Werte bis zu 30000 K. Das Strahlungsfeld ist weit vom thermodynamischen Gleichgewichtszustand entfernt, so daß der Temperaturbegriff hier völlig seinen Sinn verliert.

Das *interstellare Magnetfeld* bewirkt neben der Ausrichtung der interstellaren Staubteilchen und der damit verbundenen Polarisation des Sternlichts auch eine Zeemanaufspaltung der Spektrallinien, eine Drehung der Polarisationsebene polarisierter Radiostrahlung im interstellaren Raum und die Synchrotronemission des galaktischen Radiokontinuums. Letztere entsteht dadurch, daß die geladenen Teilchen der kosmischen Strahlung durch das interstellare Magnetfeld auf Spiralbahnen um die magnetischen Feldlinien gezwungen werden und mit der Gyrationsfrequenz zu strahlen beginnen. Sowohl die Zeemanaufspaltung der 21-cm-Linie als auch Messungen des Faradayeffekts (↑ 9.1.5.2.3.) legen für das interstellare Magnet-

feld eine mittlere Flußdichte von $3 \cdot 10^{-10}$ T nahe. Die Synchrotronstrahlung deutet auf einen um eine Größenordnung höheren Wert hin. Da wahrscheinlich ein großer Teil der beobachteten Synchrotronemission diskreten Quellen wie Supernovaüberresten entstammt, bedeutet dies keinen Widerspruch zu den anderen Messungen. Die Feldstruktur scheint kompliziert zu sein, es wird aber vermutet, daß das großräumige Magnetfeld längs der Spiralarme ausgerichtet ist.

Die Energiedichte der *kosmischen Strahlung* ist vergleichbar mit den Energiedichten des Strahlungsfelds und des magnetischen Felds. Eine solche Gleichverteilung ist nur durch eine enge Kopplung der beteiligten Felder erklärbar. Die kosmische Strahlung besteht hauptsächlich aus Protonen, Heliumkernen und den Atomkernen der schwereren Elemente. Bemerkenswert ist eine gewisse Überhäufigkeit der Lithium-, Bor-, Beryllium- und He-3-Kerne, die wahrscheinlich durch Spaltung schwerer Teilchen bei Stößen mit den Partikeln der interstellaren Materie entstehen. Die kosmische Strahlung ist nahezu isotrop und besteht hauptsächlich aus niederenergetischen Teilchen mit Energien bis zu 10^{10} eV; Teilchen höherer Energie, bis 10^{20} eV, sind sehr selten. Quellen der kosmischen Strahlung dürften die Supernovaausbrüche und die Pulsare sein. Inwieweit eruptive Vorgänge in den Sternatmosphären zur niederenergetischen Teilchenstrahlung beitragen, wie das bei der Sonne im geringen Maße der Fall ist, ist nicht bekannt.

10.3.5. Verteilung der Materie im Milchstraßensystem

Die Sonne bildet zusammen mit weiteren 10^{11} Sternen sowie großen Mengen interstellarer

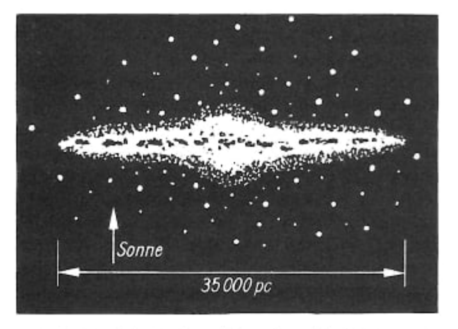


Abb. 10.3.5-1 Schematisierter Anblick des Milchstraßensystems für einen außergalaktischen Beobachter; Blick in Richtung der galaktischen Ebene; *große Punkte* stellen Kugelhaufen, *kleine Punkte* RR Lyrae-Sterne dar

Materie ein räumliches System, das durch die gravitative Anziehung seiner Bestandteile zusammengehalten wird und einer diskusförmigen Scheibe mit einer zentralen Verdichtung ähnlich sieht (Abb. 10.3.5-1). Es wird als Milchstraßensystem bezeichnet, weil die von der Erde aus gesehenen, mit bloßem Auge aber nicht mehr als einzelne Sterne erkennbaren Mitglieder des Systems die Milchstraßenerscheinung am Himmel hervorrufen. Die Erforschung der inneren Struktur des Milchstraßensystems oder der *Galaxis* wird dadurch erschwert, daß sich die Erde mitten in ihm befindet und nur rund 15% des Systems mit optischen Mitteln überblickbar sind; der Rest ist infolge der interstellaren Extinktion unsichtbar. Die Kenntnisse über den Bau der Galaxis sind daher noch unsicher, doch zeigen sich keine wesentlichen Unterschiede in bezug auf andere, extragalaktische Sternsysteme.

10.3.5.1. Populationen. Fast ausschließlich in unmittelbarer Nähe der Symmetrieebene des Systems findet man die interstellare Materie und sehr junge Sterne der Population I, z. B. die Assoziationen, offene Sternhaufen, O-Sterne, δ-Cephei-Sterne. Die Sterne der Spektralklasse A, die noch zur Population I gehören, aber älter als die O-Sterne sind, haben im Mittel einen größeren Abstand von der galaktischen Ebene; sie bilden ein etwas größeres Untersystem als die O-Sterne. Ein noch größeres Untersystem wird von den planetarischen Nebeln, den Novae und den RR Lyraesternen kürzerer Periode gebildet sowie von den meisten Sternen der Spektralklassen F bis M. Da sie zahlenmäßig am häufigsten sind, repräsentieren sie die diskusähnliche Scheibe, sie bilden die *Scheibenpopulation.* Auch die Sterne im Kern der Galaxis gehören dazu. Das räumlich ausgedehnteste, fast kugelförmige Untersystem wird von den *Kugelsternhaufen* und den RR-Lyrae-Sternen längerer Periode gebildet, d. h. von der extremen Population II. Es sind dies die ältesten Objekte im Milchstraßensystem, deren relativer Gehalt an schweren Elementen am kleinsten ist. Die einzelnen Untersysteme durchdringen sich, so daß man in Nähe der galaktischen Ebene Mitglieder aller Populationen finden kann.

Aus Sternzählungen läßt sich die Dichteverteilung der Sterne bestimmen; es lassen sich Richtungen feststellen, in denen die *Sterndichte* höher als in der Umgebung der Sonne ist. In Richtung auf die Sternbilder Scutum, Sagittarius, Scorpius treten Sterne aller Sterngruppen gehäuft auf, während in der Gegenrichtung auffällig wenige zu finden sind. Die großen Sternzahlen werden dadurch verursacht, daß in dieser Richtung der *Kern der Galaxis* liegt, in dem die räumliche Sterndichte besonders hoch ist. Wegen Dunkelwolken zwischen ihm und der Sonne ist der Kern aber optisch nicht zu erreichen. Der

Verlauf der Linien gleicher Sterndichte in einer Ebene senkrecht zur galaktischen Ebene (Abb. 10.3.5-2) zeigt die Symmetrie des Systems, die Verdickung in Richtung zum Zentrum, aber auch Deformationen der Linien gleicher Dichte in der Nähe der galaktischen Ebene sind erkennbar. Ursache der Deformationen können vorhandene Spiralarme sein.

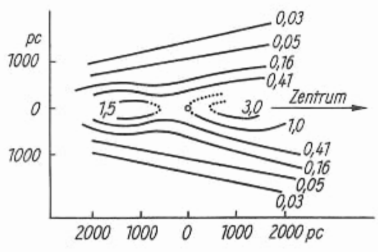


Abb. 10.3.5-2 Linien gleicher Sterndichte in einem Schnitt senkrecht zur galaktischen Ebene durch Sonne (o) und galaktisches Zentrum; die Zahlen geben die Sterndichte in Einheiten der Dichte in der Sonnenumgebung an, Abstände von der Sonne in pc

Die *Dicke* des eigentlichen *Milchstraßensystems*, das durch die Mitglieder der Scheibenpopulation und der Population I gebildet wird, beträgt im Kern etwa 5000 pc, der *Durchmesser* in der galaktischen Ebene etwa 35000 pc. Das Untersystem der Population II, das die eigentliche Galaxis in Form eines *Halos* umgibt, hat einen Durchmesser von etwa 50000 pc. Die Sonne ist etwa 10000 pc vom Zentrum entfernt und steht etwa 15 pc nördlich von der Symmetrieebene.

10.3.5.2. Bewegungsverhältnisse. Die Sterne, Sternhaufen und interstellaren Wolken bewegen sich auf Rosettenbahnen um das galaktische Zentrum, wobei die Winkelgeschwindigkeit nach außen abnimmt; die näher zum Zentrum gelegenen Sterne umlaufen es schneller als die weiter entfernten. Aus dieser *differentiellen Rotation* der Galaxis kann ihre Gesamtmasse zu rund $2 \cdot 10^{11}$ Sonnenmassen abgeschätzt werden. Der systematischen Bewegung um das Milchstraßenzentrum sind die individuellen oder *Pekuliarbewegungen* einzelner Sterne und Sterngruppen überlagert. Die Pekuliarbewegung der Sonne spiegelt sich in dem Bewegungsverhalten der Sterne der Sonnenumgebung wider. Unterwirft man die Eigenbewegungen und bekannten Radialgeschwindigkeiten aller Sterne bis zur scheinbaren Helligkeit 12^m einer statistischen Analyse, wobei vorausgesetzt wird, daß sich diese Sterne wahllos im Raum bewegen, stellt man fest, daß sich die Sonne mit einer Geschwindigkeit von 19,5 kms^{-1} auf einen Punkt, den *Apex*, im Sternbild Herkules zubewegt. Andere Geschwindigkeiten und Zielpunkte ergeben sich, wenn für eine solche Analyse Sterne, die hinsichtlich ihrer Populationszugehörigkeit oder anderer physikalischer Merkmale zusammengehören, verwendet werden. Beispielsweise ist die Relativgeschwindigkeit der Sonne gegenüber den Sternen früher Spektraltypen merklich geringer, erreicht aber bezüglich der Sterne des galaktischen Halos Werte von 200 kms^{-1}. Diese bleiben, weil sie kaum an der Rotation um das Milchstraßenzentrum teilnehmen, hinter der Sonne zurück.

Da die kinematischen Besonderheiten der Populationen offensichtlich die Bewegungsverhältnisse zum Zeitpunkt ihrer Entstehung „konserviert" haben, erlauben die Untersuchungen der Bewegungsvorgänge Einblicke in die Vergangenheit unserer Galaxis.

10.3.5.3. Spiralstruktur. Wie in den extragalaktischen Spiralsystemen kann auch in unserer Galaxis die Spiralstruktur zumindest in der Sonnenumgebung durch die Anordnung der sehr jungen Sterne und leuchtenden Gasnebel längs dieser Arme beobachtet werden (Abb. 10.3.5-3). Auf eine großräumige Spiralstruktur unseres Sternsystems weisen die 21-cm-Linienemissionen des neutralen interstellaren Wasserstoffs hin, der in den Spiralarmen konzentriert ist (Tafel 61). Da Beobachtungen in diesem Wellenlängenbereich nicht durch die interstellare Extinktion beeinträchtigt werden, können mittels der 21-cm-Linie sogar Spiralstrukturen

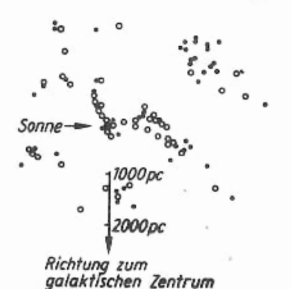


Abb. 10.3.5-3 Spiralstruktur in Sonnennähe; *Punkte* bedeuten beobachtete junge offene Sternhaufen, *Kreise* beobachtete H II-Gebiete

jenseits des Milchstraßenzentrums festgestellt werden: Die Entfernung der beobachteten HI-Wolken ergibt sich aus den gemessenen Dopplergeschwindigkeiten unter Zugrundelegung eines ebenfalls aus 21-cm-Beobachtungen ableitbaren Rotationsgesetzes für die Milchstraße. Von den Pekuliargeschwindigkeiten der Wasserstoffwolken von größenordnungsmäßig 10 kms^{-1} wird dabei abgesehen. Die optischen und radioastronomischen Untersuchungen ergeben folgendes Bild: Die Sonne befindet sich nahe der Innenseite des *Orionarms*, so benannt

nach den in diesem Arm gefundenen O-Assoziationen, die am Himmel im Sternbild Orion stehen. Der äußere oder *Perseusarm* ist etwa 2 kpc von uns entfernt. Ein weiterer Spiralarm erstreckt sich zwischen uns und dem galaktischen Zentrum und wird als *Sagittariusarm* bezeichnet. Vielfach wird der Orionarm als kurzes Verbindungsstück zwischen diesen beiden regulären Spiralarmen angesehen. Die Arme liegen nicht immer genau in der galaktischen Ebene, ihre Enden ragen zum Teil beträchtlich darüber hinaus. Aus dem Rotationssinn und der Krümmung der Arme geht hervor, daß die Galaxis ihre Spiralarme hinter sich herzieht. Ein extragalaktischer Beobachter würde unsere Milchstraße den Sb-Systemen zuordnen (↑ 10.4.1.).

Die heute weithin akzeptierte *Lintheorie* sieht in den beobachteten Spiralmustern ein Wellenphänomen und umgeht damit das Problem, einen Mechanismus zu finden, der die Verwendung der Arme infolge der differentiellen Rotation verhindert. Gemäß dieser Vorstellung breitet sich im Sternengas – wenn wir einmal die Sterne unseres Milchstraßensystems mit den Partikeln eines Gases vergleichen – eine *Dichtewelle* aus, die im Bereich der Arme für eine höhere Konzentration der Sterne als im Zwischenarmgebiet sorgt; das Sternengas wird also periodisch komprimiert. Verbunden mit einer solchen spiralförmigen Dichtewelle ist eine kleine Störung im Gravitationspotential der galaktischen Scheibe, die gerade so bemessen ist, daß sie die Verdichtung des Sternengases in den Spiralarmen reproduziert. Die Theorie zeigt, daß solche selbstkonsistente Dichtewellen in einer differentiell rotierenden Scheibe möglich sind und bevorzugt zweiarmige Spiralmuster entstehen sollten, was in Einklang mit den Beobachtungen an extragalaktischen Spiralnebeln steht. Die Amplitude der Dichtewellen kann durchaus gering sein, da bereits eine geringfügige Störung im Gravitationspotential der galaktischen Scheibe ausreicht, das interstellare Gas in den Potentialsenken zu sammeln. Der Grund dafür ist in den im Vergleich zu den Sternen wesentlich niedrigeren Pekuliargeschwindigkeiten der Wasserstoffwolken zu suchen, d. h., daß das interstellare Gas weniger turbulent ist als das „heiße Sternengas".

Da das interstellare Gas mit Überschallgeschwindigkeit gegen die Gasarme anläuft, bilden sich großräumige Stoßfronten an den Innenkanten der Spiralarme aus, die eine zusätzliche Verdichtung des Gases in diesem Bereich bewirken und so die Entstehung von Sternen begünstigen, deren massereichen Vertreter für kurze Zeit infolge ihrer großen Leuchtkraft und den sie umgebenden HII-Regionen das optische Erscheinungsbild der Spiralstruktur bestimmen.

Durch welche Prozesse Dichtewellen angeregt werden, ob durch innere, vom Galaxienkern ausgehende Einflüsse oder äußere Gezeitenwirkungen – z. B. durch den nahen Vorübergang eines anderen Sternsystems ausgelöst – ist nicht bekannt.

10.3.5.4. Galaktisches Zentralgebiet. Die Beobachtungen des galaktischen Zentrums beschränken sich auf den Radiofrequenz- und Infrarotbereich, weil die optische Strahlung aus diesem Gebiet die vorgelagerten Staubwolken nicht durchdringen kann. Die Radiostruktur ist kompliziert. Das Milchstraßenzentrum fällt mit der thermischen Komponente der Radioquelle *Sagittarius A* (Sgr A West), einer riesigen HII-Region von etwa 500 Sonnenmassen, zusammen. In der Nähe dieser auch im infraroten Spektralbereich auffälligen Quelle befindet sich nebst schwächeren thermischen Quellen auch eine starke nichtthermische Radioquelle, Sagittarius A Ost, die in ihren Ausmaßen mit einem galaktischen Supernovaüberrest vergleichbar ist. Einige kühle *Molekülwolken* vervollständigen das im Radiobereich zu beobachtende Bild. Das gesteigerte Auflösungsvermögen der radioastronomischen Empfangsapparaturen ermöglichte vor einigen Jahren die Entdeckung einer weiteren nichtthermischen Radioquelle noch innerhalb von Sgr A West. Diese extrem kompakte Komponente hat einen Durchmesser von nur 140 AE, ist zeitlich veränderlich und wird allgemein als das eigentliche Zentrum unserer Milchstraße angesehen. Das ganze Zentralgebiet ist in eine ausgedehnte, schnell rotierende Wasserstoffscheibe eingebettet.

Die vom Kerngebiet ausgehende *Infrarotstrahlung* entspricht der Ausstrahlung eines alten, sehr dichten Sternhaufens von 10^8 Sternen. Eine solche Sternkonzentration im galaktischen Zentrum – die Sterndichte übersteigt die mittlere Sterndichte in der Sonnenumgebung um das $10 \cdot 10^6$fache – wird auch durch die Beobachtung anderer Galaxienkerne nahegelegt. Eine ausgedehnte Infrarotquelle um das galaktische Kerngebiet herum wird mit der Ausstrahlung aufgeheizter Staubteilchen in Verbindung gebracht.

Die kompakte, nichtthermische Kernquelle steht vermutlich in Zusammenhang mit Aktivitäten, wie sie, allerdings wesentlich stärker, in den Kernen *aktiver Galaxien* (↑ 10.4.2.1.) gefunden werden. Die im Bereich der 21-cm-Linie beobachteten *expandierenden Wasserstoffarme* in Zentrumsnähe sind möglicherweise auf eine vor mehreren Mill. Jahren erfolgte Explosion im Zentrum unseres Milchstraßensystems zurückzuführen, bei der in mehreren Schüben vom Kern einige Mill. Sonnenmassen Wasserstoffgas ausgeworfen wurden.

10.3.5.5. Alter und Entwicklung des Milchstraßensystems. Die ältesten Sterne in unserer Milchstraße gehören der Halopopulation an

und dürften 10 bis 15 Mrd. Jahre alt sein. Man nimmt an, daß unsere Galaxis aus einer kollabierenden prägalaktischen Gaswolke, die langsam rotierte, entstand. Durch Fragmentierung bildeten sich zunächst die Kugelsternhaufen und andere Mitglieder der extremen Population II (Halopopulation). Diese erste Phase nahm nur relativ wenig Zeit, rund 10^8 Jahre, in Anspruch; die entstandenen Sterne und Kugelsternhaufen bewegen sich auf langgestreckten Bahnen um das Schwerezentrum und zeichnen sich durch einen relativ niedrigen Gehalt an schweren Elementen aus.
Das verbleibende Gas sammelte sich sodann in einer schnell rotierenden, abgeflachten Scheibe an, in der der Prozeß der Sternentstehung auch heute noch anhält. Die chemische Zusammensetzung des interstellaren Gases ändert sich durch die allmähliche Anreicherung mit schweren Elementen, die im Inneren massereicher Sterne produziert wurden und bei Supernovaexplosionen in großen Mengen freigesetzt werden. Die erst jüngst entstandenen Sterne sind infolgedessen „metallreicher" als die alten Sterne der Halopopulation oder die ersten in der galaktischen Scheibe entstandenen Sterne.
Die Deutung der beobachteten Elementehäufigkeiten in Sternen unterschiedlicher Populationszugehörigkeit gestattet deshalb, die chemische Entwicklung unseres Milchstraßensystems von den Anfängen an zu verfolgen. Von einer vollständigen Lösung der mit dem galaktischen Evolutionsgeschehen verbundenen Probleme sind wir heute allerdings noch weit entfernt.

10.4. Erscheinungsformen der Materie in extragalaktischen Sternsystemen

Am Himmel sind schwache, nebelhafte Gebilde zu erkennen, der Andromedanebel und die beiden Magellanschen Wolken am Südhimmel sogar mit bloßem Auge. In etwa 100 dieser Nebel konnte der amerikanische Astronom Hubble Kugelhaufen, offene Sternhaufen, Novae, Supernovae, Veränderliche, Überriesen, leuchtende Gasnebel und dunkle staubförmige Materie nachweisen und ihre Entfernungen bestimmen. Es zeigte sich, daß es Sternsysteme außerhalb der Milchstraße sind, extragalaktische Sternsysteme, auch *Galaxien* [Singular Galaxis oder Galaxie] genannt. Einige helle Galaxien werden nach dem Sternbild bezeichnet, in dem man sie sieht, z. B. Andromedanebel, Triangulumnebel oder mit der Nummer eines Nebelkatalogs; dabei bedeuten M Messier-Katalog von 1784, NGC New General Catalogue of Nebulae and Clusters of Stars von 1888 und IC Index Catalogue, Ergänzungen zum Allgemeinen Neuen Katalog der Nebel- und Sternhaufen. Diese großen Kataloge enthalten nur von ausgewählten Gebieten schwächere Objekte als 13^m.

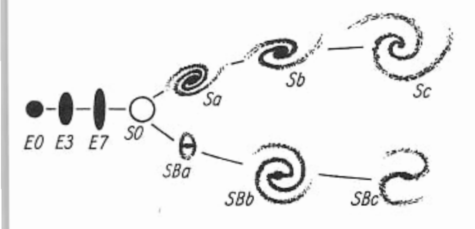


Abb. 10.4.1-1 Morphologische Klassifikation der extragalaktischen Sternsysteme nach Hubble; elliptische Nebel E0 bis E7, S0-Galaxien, normale Spiralnebel Sa, Sb, Sc, Balkenspiralen SBa, SBb, SBc

10.4.1. Typen der Galaxien

Hubble unterschied fünf Typen von Galaxien: elliptische Nebel E, S0-Galaxien, Spiralnebel S, Balkenspiralen SB und irreguläre Nebel I (Abb. 10.4.1-1; Tafel 61). *Elliptische Nebel* zeigen nur selten Strukturen; ihre Helligkeit nimmt von einem kleinen hellen Kern gleichmäßig nach außen ab. Nach dem Grad der Elliptizität werden die elliptischen Nebel in acht Unterklassen eingeteilt. Bezeichnet a die große und b die kleine Halbachse der beobachteten Nebelellipse, so kennzeichnet die auf ganze Zahlen gerundete Größe $10(a-b)/a$ den Grad der Elliptizität, z. B. bedeutet E 7, daß $b = 0{,}3$, wenn $a = 1$, und E0 einen kreisförmigen Nebel wegen $a = b$. Nach statistischen Untersuchungen kommen wahre Abplattungen zwischen 0 und 7 vor; am häufigsten ist jedoch die Abplattung 4 vertreten. Bei geeigneter Blickrichtung kann aber selbst ein stark abgeplattetes System völlig rund erscheinen. Die Sterne in den elliptischen Galaxien gehören der Population II an, sind mithin sehr alte Sterne. Interstellares Gas ist in diesen Galaxien nur sehr selten enthalten.
Spiralnebel S zeigen Spiralarme, wenn man mehr oder weniger auf die Hauptebene sieht. In Blickrichtung auf die Kante erscheinen sie spindelförmig, mit dunklen Bändern von stark absorbierender Materie durchsetzt. Die *Spiralarme* entspringen normalerweise tangential am Rande eines sehr hellen Kerngebiets an gegenüberliegenden Punkten und winden sich über etwas mehr als einen vollen Umlauf in Form einer logarithmischen Spirale um das Zentralgebiet. Oft sind nur zwei Arme vorhanden; meist findet man jedoch noch kleine, sekundäre Arme, die die regelmäßige Form stark verwischen können. Spiralnebel Sa haben ein sehr großes, helles Kerngebiet; bei den Typen Sb ist der Kern unauffälliger und bei Sc ist er nur noch schwach

erkennbar als sehr kleines, rundes Gebilde. Die Spiralarme winden sich bei dem Typ Sa eng um das Kerngebiet, öffnen sich bei den Sb- bis Sc-Spiralen immer weiter und werden dabei breiter und verwaschener. In neuerer Zeit unterscheidet man noch zwei Untergruppen Sd und Sm mit stark aufgelockerten Armen und sehr kleinem Kern, die als Übergangsformen zu den irregulären Galaxien gelten.

In den Spiralarmen findet man vorwiegend Objekte der Population I: sehr helle, junge Sterne vom Spektraltyp O und B, leuchtende Gasnebel und dunkle absorbierende Materie. Das Kerngebiet besteht aus Vertretern der Population II. Der ganze Spiralnebel ist eingebettet in ein etwas größeres, sphärisches System, das ähnlich wie der Halo der Milchstraße bei den näheren Galaxien an zahlreichen Kugelhaufen erkennbar ist; beim Andromedanebel sind es rund 200.

Balkenspiralen SB unterscheiden sich von den gewöhnlichen Spiralgalaxien durch einen „Balken", der den Kern geradlinig durchzieht. Von beiden Enden entspringen die Spiralarme bei den Typen SBa und SBb rechtwinklig und umwinden den Kern fast kreisförmig. Die SBc-Typen haben eine S-förmige Gestalt; die Arme sind weit geöffnet. Analog wie bei den normalen Spiralnebeln ist die Größe des Kerngebiets ein Klassifizierungsmerkmal für die Untergruppen a bis m.

Der *SO-Typ* gleicht dem Kerngebiet der Typen Sa bzw. SBa und besteht aus einer nahezu gleichmäßig hellen Scheibe oder Linse mit scharfem Rand und einem sehr hellen punktförmigen Kern. Spiralarme sind nicht vorhanden, aber in etwas größerem Abstand enthält dieser Typ gelegentlich einen schwach leuchtenden Ring, der sich nach außen allmählich verliert.

Die *unregelmäßigen* oder *irregulären Galaxien* I haben keine Symmetrieeigenschaften und keinen Kern. Ihre regellose Struktur besteht aus hellen Kondensationen und Gebieten dunkler Staubmaterie. Sie bestehen aus Objekten der Population I.

Neben diesen Hubbletypen unterscheidet man

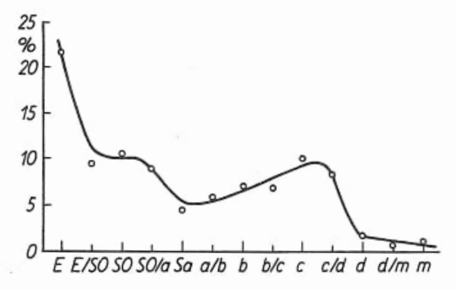


Abb. 10.4.1-2 Häufigkeitsverteilung der Untergruppen nach DE VAUCOULEURS

in den letzten Jahren noch weitere besondere Galaxienformen nach ihrem Aussehen: *Seyfertgalaxien* besitzen ein abnorm helles Kerngebiet, können aber nach ihrem sonstigen Aussehen zu einem der Hubbletypen zugeordnet werden. *N-Galaxien* sind sehr kompakt, meist strukturlos, und besitzen eine hohe Flächenhelligkeit. *Quasistellare Objekte* QSO erscheinen sternförmig und sind nur durch ihre Rotverschiebung (↑ 10.4.2.1.) als extragalaktische Sternsysteme zu erkennen. Ein Merkmal dieser Objekte ist ihr großer UV-Exzeß, d. h. ihre abnorme große Helligkeit im blauen und ultravioletten Farbbereich im Vergleich zu normalen Sternen. Dieses Merkmal weisen auch die sogenannten *Markariangalaxien* auf.

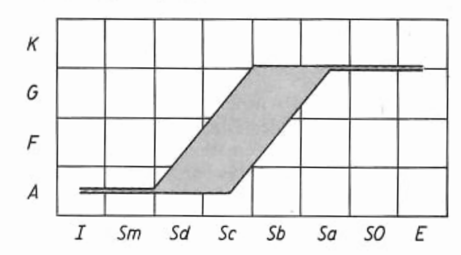


Abb. 10.4.2-1 Spektrum der Nebeltypen nach DE VAUCOULEURS

Die *Häufigkeit* der morphologischen Haupttypen der Galaxien beträgt für E 13 %, für SO 21,5 %, für S' 61,1 % und für I 3,4 %, d. h., etwa 60 % der beobachteten Galaxien sind Spiralnebel, wobei normale Spiralen und Balkenspiralen gleich häufig sind. Man faßt sie deshalb oft in Häufigkeitsverteilungen nach Untergruppen zusammen (Abb. 10.4.1-2). Die „späten" Typen Sd und Sm sind wahrscheinlich häufiger, da nicht alle wegen ihrer geringen Leuchtkraft beobachtet werden können. Dagegen sind die QSO seltene Objekte; man schätzt 1 auf 10^5 bis 10^7 normale Galaxien.

10.4.2. Integrale Zustandsgrößen

Von entfernten Galaxien können Einzelheiten nicht untersucht werden; man kennt nur integrale Zustandsgrößen des Nebels als Ganzes, kann aber aus ihnen auf Leuchtkraft, Masse, Entfernung und Ausdehnung Schlüsse ziehen, wenn man diese an benachbarten Galaxien „geeicht" hat.

10.4.2.1. Spektrum und Rotverschiebung. Das Spektrum einer normalen Galaxie ist zusammengesetzt aus den Spektren aller Einzelsterne und der leuchtenden interstellaren Materie und variiert zwischen einem „frühen" F- bzw. A-Typ und dem „späten" G-Typ (Abb. 10.4.2-1).

Elliptische Nebel enthalten Sterne der Population II, d. h. rote Riesen und Zwerge, und erscheinen daher röter als Spiralnebel und irreguläre Galaxien, die außerhalb des Kerngebiets eine große Anzahl heißer blauer Objekte der Population I enthalten. Die Kerne der S-Systeme haben ebenfalls einen späteren Spektraltyp als die Randpartien. In *aktiven Galaxien*, insbesondere bei den quasistellaren Objekten, besteht auch das optische Kontinuum nicht aus einer thermischen Strahlung, sondern ist *Synchrotronstrahlung* (↑ 10.4.2.4.).

Starke *Emissionslinien* von Wasserstoff, Sauerstoff und Neon sind auch aus den Spektren galaktischer Emissionsnebel bekannt; ihr Vorkommen weist auf ausgedehnte Emissionsgebiete in der Nähe heißer Übergiganten vom Typ O und B hin. Im Einklang damit steht, daß Emissionslinien bei den „späten" Galaxien vom Typ S oder I wesentlich häufiger sind als bei den elliptischen oder SO-Systemen. Auch die 21-cm-Linie des neutralen Wasserstoffs wird in Spiralsystemen beobachtet. Sehr breite Emissionslinien treten in Seyfertgalaxien und in einigen elliptischen Riesengalaxien auf. Sie deuten auf große *turbulente Bewegungen* des interstellaren Gases mit Geschwindigkeiten bis 1000 kms^{-1} hin.

Wegen der großen Entfernung können an den Galaxien keine Eigenbewegungen nachgewiesen werden. Man war auf die Messung der Radialgeschwindigkeit angewiesen, die sich nach dem Dopplereffekt aus der Linienverschiebung im Spektrum der Lichtquelle bestimmen läßt. Hierbei fand man, daß der überwiegende Teil der Galaxien eine Verschiebung $z = \Delta\lambda/\lambda$ des Spektrums nach größeren Wellenlängen λ, also nach dem Roten, zeigt. Nach der speziellen Relativitätstheorie ergibt sich aus den gemessenen z-Werten als *Fluchtgeschwindigkeit* $V = c[(1+z)^2 - 1]/[(1+z)^2 + 1]$, die für kleine z-Werte ($z \leqq 0{,}2$) in die klassische Dopplerformel $V = cz$ übergeht, für $z = 2$ aber $V = 0{,}8$ c, für $z = 5$ den Wert $V = 0{,}95\,c$ ergibt und für $z \to \infty$ gegen $V = c$ geht. Der Betrag der *Rotverschiebung* ist um so größer, je kleiner der scheinbare Durchmesser bzw. die scheinbare Helligkeit der Galaxien ist (Abb. 10.4.2-2), je größer also ihre Entfernung ist; man deutet diesen von HUBBLE entdeckten Effekt als *Expansion der Welt*. Diese Ausdehnung des heute überschaubaren Weltalls zeichnet keine Galaxie aus; von jeder anderen aus würde ein Beobachter die gleiche Erscheinung feststellen.

Es besteht im bisher bekannten Bereich als linearer Zusammenhang die *Hubblebeziehung* $V = H \cdot r$ zwischen Entfernung r und Fluchtgeschwindigkeit V. Hierin ist H die *Hubblekonstante*, deren Wert 50 kms^{-1} Mpc^{-1} mit einer Unsicherheit von 10% beträgt.

Die Hubblebeziehung wird benutzt, um für weitentfernte Sternsysteme aus der beobachteten Rotverschiebung die Entfernung zu bestimmen. Die Deutung des Hubbleeffekts muß im Rahmen der Kosmologie erfolgen. Diese allgemeine Fluchtbewegung wird statistisch überlagert von *Pekuliarbewegungen* von 50 bis 150 kms^{-1} und in Galaxienhaufen bis über 1000 kms^{-1}. In den Galaxien der näheren Umgebung kommen dadurch auch Bewegungen auf uns zu vor.

10.4.2.2. Entfernung und Leuchtkraft. Aus der scheinbaren Helligkeit m läßt sich die Entfernung r nach $\lg r = 1 + 0{,}2(\mathrm{m} - \mathrm{M})$ finden, wenn die absolute Helligkeit M bekannt ist. Diese fotometrische Methode läßt sich bei nahen Sternsystemen auf Einzelobjekte anwenden. Aus den Perioden der δ-Cephei-Sterne ergibt sich ihre absolute Helligkeit von -1^m bis 5^m (↑ 10.3.3.1.). Auch aus den bekannten absoluten Helligkeiten der O- und B-Überriesen, der Novae und der Supernovae sowie der Kugelhaufen läßt sich mit dieser fotometrischen Methode die Entfernung von Galaxien bestimmen. Außerdem läßt sich aus dem Vergleich des scheinbaren Durchmessers d in Bogenminuten von Wasserstoffemissionsgebieten, HII-Gebieten, mit dem wahren Durchmesser D in pc solcher Gebiete, der allerdings vom morphologischen Typ der Galaxien abhängt, die Entfernung in kpc nach der Beziehung $r = 3{,}44\ D/d$ berechnen.

Die *absoluten Helligkeiten* der elliptischen Galaxien streuen über den weiten Bereich von -9^m bis -22^m. Die hellsten elliptischen Nebel in einem Haufen haben im Mittel eine Helligkeit von -21^m. Die absoluten Helligkeiten der Spiralsysteme liegen zwischen -16^m und -22^m.

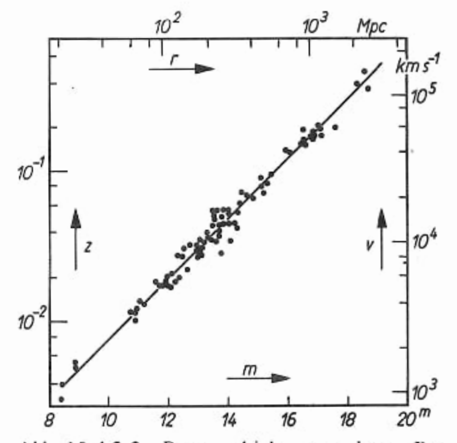


Abb. 10.4.2-2 Rotverschiebung z bzw. Radialgeschwindigkeit v in kms^{-1} als Funktion der scheinbaren Helligkeit m bzw. der Entfernung r in Mpc für die hellsten Galaxien in 82 Haufen nach SANDAGE

Unser Milchstraßensystem mit -19^m und der Andromedanebel mit -20^m zählen zu den absolut hellen Galaxien.

Der *wahre Durchmesser* der Galaxien reicht von unterhalb 0,1 kpc bei den Zwergsystemen zu über 50 kpc bei den Riesengalaxien.

10.4.2.3. Dynamik, Masse. Die vorherrschende Bewegung ist die *Rotation* der Galaxie um das Zentrum. Sie läßt sich spektroskopisch aus der Radialgeschwindigkeit einzelner Punkte in verschiedenen Abständen vom Zentrum bestimmen. Bei Spiralsystemen erfolgt die Bewegung so, daß die konvexe Seite der Arme in Bewegungsrichtung vorne ist. Bei sphärischen elliptischen Nebeln rotiert das System nicht als Ganzes, sondern die Sterne vollführen selbständige Bahnen im Gravitationsfeld der Galaxie um das Zentrum, wobei die Bahnebenen gleichmäßig im Raum verteilt sind. Eine Abplattung weist darauf hin, daß das System eine mehr oder weniger starke geordnete Rotationsbewegung besitzt. Da die Umlaufperiode eines Himmelskörpers um eine Zentralmasse nach dem Gravitationsgesetz eng mit der Größe dieser Masse verknüpft ist, kann aus der Rotationsbewegung in den Galaxien ihre Masse bestimmt werden. Weitere Massebestimmungen ergeben sich dann, wenn zwei Galaxien gravitativ gebunden sind und um einen gemeinsamen Schwerpunkt kreisen. Schließlich ist in Galaxienhaufen die Geschwindigkeitsdispersion der einzelnen Mitglieder mit der Gesamtmasse des Nebelhaufens nach dem sogenannten Virialsatz verknüpft, wenn der Haufen stabil ist. Daraus läßt sich die Masse der Galaxien bestimmen.

Die Beobachtungen haben ergeben, daß die inneren Partien der Galaxien wie ein starrer Körper rotieren, d. h. mit konstanter Winkelgeschwindigkeit. Die Umlaufgeschwindigkeit nimmt daher in diesem Gebiet mit wachsendem Abstand vom Zentrum linear zu; ein voller Umlauf dauert etwa 10 bis 100 Mill. Jahre. Die Rotationsgeschwindigkeit erreicht in einem gewissen Abstand vom Zentrum ein Maximum und fällt dann allmählich ab, um schließlich in die vom Planetensystem her bekannte Keplerbewegung überzugehen. Aus der Rotation erhält man Massen zwischen $10^8 M_\odot$ und $10^{11} M_\odot$ (Sonnenmasse $M_\odot = 1{,}99 \cdot 10^{30}$ kg), elliptische Riesensysteme erreichen $10^{13} M_\odot$ (Tab. 10.4.2-3).

Für die Bestimmung der Massendichte im Kosmos reichen die genauen Massebestimmungen an einigen Einzelgalaxien nicht aus. Hier hilft man sich, in dem man das *Masse-Leuchtkraft-Verhältnis* ermittelt. Es beträgt in Einheiten der Sonnenwerte bei Spiralsystemen und irregulären Galaxien 6,5 und bei elliptischen Systemen 30. Aus der scheinbaren Helligkeit läßt sich über die Entfernung die Leuchtkraft sehr leicht ermitteln und damit bei bekanntem morphologischem Typ aus dem Masse-Leuchtkraft-Verhältnis die Masse des Systems errechnen.

10.4.2.4. Radiogalaxien. Normale Galaxien, wie unsere Milchstraße und der Andromedanebel, senden nur eine relativ geringe Radiostrahlung etwa von der Stärke 10^{30} bis 10^{32} W

Tab. 10.4.2-3 Einige Galaxien heller als 9. Größe; E_p elliptischer Nebel mit Besonderheiten; l und b galaktische Länge bzw. Breite in Grad, Durchmesser D' in Bogenminuten und D_e linear in kpc, r_p und r_L Entfernung in Mpc und in 10^6 ly; m_v, M_v scheinbare und absolute Helligkeit im visuellen Spektralbereich in Größenklassen, V Radialgeschwindigkeit in km s^{-1}, (+ bedeutet, die Galaxie entfernt sich) M Masse in 10^9 Sonnenmassen $M_\odot$ ($1\ M_\odot = 1{,}99 \cdot 10^{30}$ kg)

Galaxie												
M	NGC	Typ	l	b	D'	D_e	r_p	r_L	m_v	M_v	V	M
	55	Sc	332	−76	25	12	1,9	6,2	7,1	−19,2	+180	30
	253	Sc	75	−89	22	13	2,2	7,2	7	−20	−70	12
	2403	Sc	150	+29	20	12	2	6,5	8,5	−19	+190	5
81	3031	Sb	141	+41	20	16	3,0	9,8	6,9	−20,9	+80	110
82	3034	I	141	+41	8	7	3	10	8,2	−19,6	+400	8
	3115	E 7	248	+37	4	5	4	13	9,1	−19,4	+430	80
104	4594	Sa	299	+51	6	8	4,4	14,4	8,1	−20,4	+1020	160
94	4736	Sb	124	+76	7	10	6	20	8,2	−20,6	+350	
64	4826	Sb	319	+83	8	12	6	20	8,4	−20,5	+360	
	4945	S	305	+12	12	14	5	16	7	−22		
63	5055	Sb	105	+74	10	15	4	13	8,6	−20,1	+2600	44
	5128	Ep	310	+19	14	15	3,8	12,4	6	−23	+260	
51	5194	Sc	105	+68	9	9	2	6,5	8,3	−19,6	+550	100
83	5236	Sc	316	+33	10	12	4	13	7,6	−20,9	+320	
101	5457	Sc	102	+60	20	23	3	10	8,1	−20,2	+400	230
	7793	Sd	4	−77	6	4	3	10	8,8	−18,6	+290	

aus. Sie rührt von der thermischen Strahlung des ionisierten Wasserstoffs im interstellaren Gas, von der nichtthermischen Synchrotronstrahlung der Supernovaereste und der Elektronen im Halo her sowie von der Linienstrahlung des neutralen Wasserstoffs. Der Hauptteil der Strahlung ist auf Kern und Scheibe konzentriert. Man kann daher die normale Radiostrahlung nur in den benachbarten Galaxien genau untersuchen. Dabei erscheinen diese Galaxien sehr ausgedehnt, z. B. ist das Radiobild des Andromedanebels 3- bis 4mal größer als das optische und zeigt eine nahezu sphärische Kontur im Unterschied zu der Abplattung im sichtbaren.

Eine 10^6mal stärkere Strahlung entsteht in den Radiogalaxien. Sie wurden zuerst als Radioquellen entdeckt; ihre Identifizierung mit optischen Objekten ist aber nur für einen kleinen Bruchteil gelungen, da die Reichweite der Radioteleskope größer als die der optischen Teleskope ist. So wäre die zweitstärkste Radioquelle (Cygnus A) selbst in 50mal größerer Entfernung noch als Radiostrahler zu erkennen, während man sie selbst mit den größten optischen Teleskopen nicht mehr entdecken konnte. In der ersten Zeit wurden die galaktischen und extragalaktischen Radioquellen nach dem Sternbild unter Hinzufügung eines Buchstabens bezeichnet, z. B. Virgo A; diese Radioquelle ist identisch mit der Galaxie M 87 im Messierkatalog bzw. NGC 4486 im New General Catalogue. Inzwischen ist die Anzahl der Radioquellen so stark gestiegen, daß man sie mit der Nummer in einem der Radioquellenkataloge kennzeichnet, z. B. 3C 270 die 270. Quelle im 3. Cambridgekatalog, identisch mit NGC 6421.

Die Radiogalaxien senden eine kontinuierliche Radiostrahlung aus, deren Intensität F mit zunehmender Frequenz ν abnimmt. Die Strahlung ist deshalb nicht thermisch, sondern Synchrotronstrahlung, die Elektronen großer Geschwindigkeit, relativistische Elektronen, in einem äußeren Magnetfeld aussenden. Nach dem Ansatz $F \sim \nu^{-\alpha}$ charakterisiert der Spektralindex α die Abhängigkeit der Strahlungsintensität von der Frequenz ν. Die Anzahl $N(E)$ der Elektronen mit der Energie E ist dann $N(E) \sim E^{-(1+2\alpha)}$ (Energiespektrum der Elektronen). Für positives α nehmen F und $N(E)$ mit wachsender Frequenz bzw. Energie ab. Aus $\lg F = -\alpha \lg \nu + \text{const}$ kann α aus einem $\lg F - \lg \nu$-Diagramm bestimmt werden; für konstantes α ergibt sich eine Gerade mit der Neigung α, und man spricht von einem linearen Spektrum. Erfahrungsgemäß liegt α zwischen 0,6 und 1,0 mit einer starken Häufung bei 0,8.

Interferometrische Untersuchungen mit großem Winkelauflösungsvermögen haben gezeigt, daß die *Struktur* der Radioquellen sehr mannigfaltig sein kann. Meist entsteht die Strahlung einer Radiogalaxie in zwei ausgedehnten Gebieten (Abb. 10.4.2-4; Tafel 62), die oft weit außerhalb der optisch sichtbaren Galaxie liegen, aber meist mehr oder weniger symmetrisch zu ihr angeordnet sind. Die wahren Abstände der beiden Radiokomponenten können bis zu 300 kpc und der Durchmesser der Komponenten kann 100 kpc erreichen. In einigen Fällen besteht die Radioquelle aus einer im Kern der

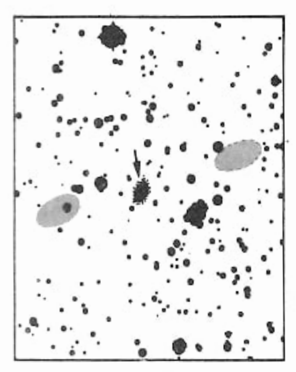


Abb. 10.4.2-4 Cygnus A (3 C 405); die beiden Emissionsgebiete der Radiostrahlung sind 200 kpc voneinander entfernt; ihr Durchmesser beträgt rund 30 kpc; in der Mitte zwischen ihnen befindet sich das durch einen *Pfeil* gekennzeichnete optische Objekt; es ist 178 Mpc von uns entfernt

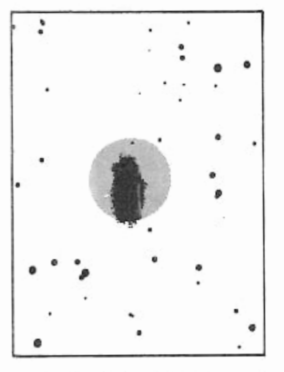


Abb. 10.4.2-5 3C 278; die Radiostrahlung kommt aus dem kreisförmigen Gebiet; ein enges Paar elliptischer Galaxien befindet sich im Zentrum (NGC 4782 und 4783)

Galaxie konzentrierten Komponente und einer ausgedehnten Halokomponente, wie bei 3C 278 (Abb. 10.4.2-5) und Virgo A. Für optische Besonderheiten von Radiogalaxien können hier nur einige Beispiele herausgegriffen werden. Bei *Virgo A* (NGC 4486, Tafel 62) ragt aus dessen

Kerngebiet ein heller Materiestrahl nach außen. Sein Licht ist stark polarisiert und läßt auf optische Synchrotronstrahlung schließen. Der Strahl besitzt Knoten hoher Radioleuchtkraft. Bei *Centaurus A* (NGC 5128, Tafel 63) ist die auffälligste Erscheinung ein sehr breiter Streifen dunkler Staubmaterie. Die beiden ausgedehnten Radioemissionsgebiete erstrecken sich bis 400000 Lichtjahre von der Galaxie entfernt. In der Galaxie selbst befindet sich noch ein enges Paar von Radioquellen. Diese sind nur 3 Bogenminuten oder rund 13000 Lichtjahre von der durch das Staubband markierten Symmetrieebene entfernt.

Aus der Strahlungsleistung läßt sich die in den relativistischen Elektronen und im Magnetfeld gespeicherte Energie abschätzen (Tab. 10.4.2-6). Man erhält Werte von 10^{50} bis 10^{53} J. Zum Vergleich sei das gesamte Masse-Energie-Äquivalent einer Riesengalaxie von $10^{12} M_0$ mit 10^{58} J angeführt. Die potentielle Energie eines solchen Systems bei einem effektiven Radius von 1 kpc beträgt 10^{55} J. Die Energieverluste durch die beobachtete Strahlung erlauben nur eine *Lebensdauer* der Radioemsssion von 10^5 bis 10^6 Jahren. Das radioeruptive Stadium kann deshalb nur ein vergleichsweise kurzes Stadium im Leben einer Galaxie sein. Für die hohen Energiebeträge wäre die Ruhemasse von 10^6 bis 10^7 Sonnen erforderlich. Man benötigt 10^8 bis 10^9 Sonnen bei einem Wirkungsgrad von 1% für die Umwandlung von Masse in Energie. Zur Zeit ist noch kein Mechanismus bekannt, nach dem diese Umwandlung vor sich gehen könnte. Das Auftreten von *Radiodoppelquellen* spricht dafür, daß sie das Ergebnis gigantischer Explosionen in den Kernen der Galaxien sind. Dabei werden ionisierte Gaswolken, Plasma, mit den eingeschlossenen Magnetfeldern und schnellen Elektronen ausgeschleudert. Bei Centaurus A hätten demnach zwei Explosionen stattgefunden: Eine vor $4 \cdot 10^5$ Jahren und eine vor $1{,}3 \cdot 10^3$ Jahren, wenn man annimmt, daß das Material nahezu mit Lichtgeschwindigkeit

Tab. 10.4.2-6 Einige Radiogalaxien; m_{pg} scheinbare fotografische Helligkeit; r Entfernung in Mpc, F beobachtete Radioflußdichte in 10^{-26} W m^{-2} Hz^{-1}, L Radioleuchtkraft in W, d Durchmesser der Radioemmisionsgebiete in kpc, A gegenseitiger Abstand der Doppelquellen in kpc; E Energie der relativistischen Teilchen in J, B Stärke des Magnetfeldes in T

Galaxie	Typ	m_{pg}	r	F	L	d	A	E	B
Centaurus A (NGC 5128)	Ep	7,87	4,0	5800					
äußeres System					$4 \cdot 10^{34}$	150	300	$8 \cdot 10^{52}$	0,05
inneres System					$1 \cdot 10^{34}$	4	10	$2 \cdot 10^{50}$	0,63
Cygnus A	–	17,9	170	8300	$2 \cdot 10^{38}$	26	200	$4 \cdot 10^{53}$	4,0
3C 295	–	20,9	1300	70	$2 \cdot 10^{38}$	28	60	$2 \cdot 10^{53}$	2,5
Virgo A (NCC 4486)	E	9,56	11	1000					
Halo					$5 \cdot 10^{34}$	50	–	$4 \cdot 10^{51}$	0,3
Zentralgebiet					$5 \cdot 10^{34}$	2,5	–	$4 \cdot 10^{50}$	2,0
Kern					$5 \cdot 10^{32}$	0,002	–	$5 \cdot 10^{43}$	25,0
M 82	I	9,20	2	13	$3 \cdot 10^{32}$	0,9	–	$1 \cdot 10^{48}$	0,79

Tab. 10.4.2-7 Einige quasistellare Objekte; z Rotverschiebung, m_v scheinbare visuelle Helligkeit, r Entfernung in Mpc, M_v absolute Helligkeit, r und M_v gelten für den Einstein-de-Sitter-Kosmos ($q_0 = {}^1/_2$)

QSO	z	m_v	r	M_v
Ton 256	0,131	15,9	780	−23,6
3C 273	0,158	12,8	950	−27,1
3C 48	0,367	16,2	1740	−25,7
CTA 102	1,037	17,3	3600	−27,1
3C 270,1	1,519	18,6	4400	−26,7
3C 191	1,953	18,4	5000	−27,5
3C 9	2,012	18,2	5100	−27,7
PKS 0237-23	2,223	16,6	5300	−29,6
OQ 172	3,53	17,9	6400	−29,4

ausgestoßen wurde. Dieser Vorgang kann bei der irregulären Galaxie M 82 (NGC 3034) direkt nachgewiesen werden. Eine Aufnahme dieses Systems zeigte starke irreguläre Strukturen (Tafel 63). Aus dem Zentrum strömen nach beiden Seiten riesige Wasserstoffwolken von mehreren Mill. Sonnenmassen und mit Geschwindigkeiten bis 10^6 ms^{-1}. Die höchste Geschwindigkeit wurde in den am weitesten vom Kern entfernten Partien (3 bis 4 kpc) festgestellt. Daraus kann man errechnen, daß die Explosion des Kerns vor $1,5 \cdot 10^6$ Jahren stattgefunden haben muß.

Der Abstand der Komponenten bei verschiedenen Radiogalaxien ist ein *Alterseffekt*; er ist um so größer, je weiter der Ausbruch zeitlich zurückliegt.

10.4.2.5. Quasistellare Radioquellen, Quasare (QSS). Im Unterschied zu Radiogalaxien haben Quasare ein quasistellares sternförmiges Aussehen. Nur ihr Radiobild zeigt ebenfalls oft Radioquellen außerhalb des optischen Objekts. Ihre außerordentlich große Rotverschiebung bis $z = 3,5$ ordnet sie den extragalaktischen Objekten zu. Nach der Hubblebeziehung sind es die entferntesten bisher bekannten Systeme (Tab. 10.4.2-7). Im optischen Spektrum treten Emissions- und Absorptionslinien auf, wie sie auch von planetarischen Nebeln (↑ 10.3.2.) bekannt sind. Aus dem Auftreten verbotener Linien, den beobachteten Linienstärken, lassen sich die *physikalischen Bedingungen* der Emissionsregion berechnen. So beträgt die Elektronendichte 10^4 bis 10^6 cm^{-3}, die Elektronentemperatur 10^4 K und der Durchmesser des Emissionsgebiets zwischen 1 und 10 pc. Absorptionslinien zeigen im gleichen Quasar oft ein System von unterschiedlichen Rotverschiebungen. Dies deutet darauf hin, daß die Absorptionslinien in Wolken entstehen, die sich bis zu $0,3\,c$ relativ zu den Emissionsgebieten bewegen und vermutlich vom *Quasarkern* ausgeschleudert wurden. Viele Quasare sind im sichtbaren Bereich veränderlich. Diese Variabilität erfaßt aber nur das Kontinuum. Aus der Zeitskala der Variabilität von Tagen bis Monaten

Tab. 10.4.3-1 Lokale Gruppe; p beim Typ gibt Besonderheiten im Aussehen an, l und b galaktische Länge bzw. Breite in Grad, D' Durchmesser in Bogenminuten und D_e linear in kpc, r Entfernung in kpc (eingeklammerte Werte sind unsicher) m_v und M_v scheinbare bzw. absolute visuelle Helligkeit in Größenklassen; V Radialgeschwindigkeit in km s^{-1}, M Masse in 10^9 Sonnenmassen M_0 ($1\,M_0 = 1,99 \cdot 10^{30}$ kg)

Galaxie	Typ	l	b	D'	D_e	r	m_v	M_v	V	M
Milchstraße	Sb	–	–	–	30	–	–	−20,3	–	100
Große Magellanwolke	I	280	−33	432	6,3	50	0,1	−18,6	+280	14
Kleine Magellanwolke	I	303	−45	216	3,1	50	2,4	−17,6	+167	2
Andromedanebel										
M31 NGC 224	Sb	121	−21	163	33	690	3,5	−21,3	−270	370
M33 NGC 598	Sc	135	−32	62	14	720	5,8	−18,9	−190	14
M32 NGC 221	E2	121	−22	3	0,7	690	8,2	−16,6	−210	4
NGC 205	E5p	121	−21	12	2,4	690	8,2	−16,6	−240	8
Sculptorsystem	E	284	−84	45	0,7	50	7	−12		0,3
Fornaxsystem	E	237	−65	50	1,6	110	7	−13	+40	1
NGC 6822	I	26	−19	20	2,3	480	8,7	−15,4	−40	0,4
NGC 147	Ep	120	−14	8	1,4	690	9,6	−15,2		1
NGC 185	Ep	121	−14	12	1,0	690	9,5	−15,3	−340	1
IC 1613	I	129	−60	15	3,0	690	9,6	−15,2	−240	0,1
Wolf-Lundmark-System	E5	74	−73	10	1,5	(500)	10,8	−13,5	−80	0,1
Leosystem I	E4	227	+49	12	0,9	260				
Leosystem II	E1	222	+68	11	0,6	180	12	−10		
Mögliche Mitglieder										
IC 10	Sc	119	−3	4					−340	
IC 342	Sc	139	+10	30					−10	
NGC 6946	Sc	97	+11	7		(800)	9	−17	+40	
Leosystem III	I	197	+54	3,0		(1000)	13			
Sextanssystem	I	246	+40				11		+370	
NGC 300	Sc	299	−80	20		(1000)	8,5	−16	+250	

läßt sich schließen, daß das Emissionsgebiet für die Kontinuumstrahlung (es ist vorwiegend Synchrotronstrahlung) sehr klein sein muß (10^{15} bis 10^{18} cm).
Da sich die Quasare durch ihren hohen blauen und ultravioletten Anteil im Spektrum von den normalen Sternen unterscheiden, führte die optische Suche nach solchen blauen Objekten zur Entdeckung der *blauen kompakten Galaxien* oder *quasistellaren Galaxien* (QSG). Sie unterscheiden sich von den Quasaren nur durch das Fehlen einer über der Nachweisgrenze liegenden Radiostrahlung. Man faßt daher die QSS und die QSG zum einheitlichen Typ der *quasistellaren Objekte* (QSO) zusammen.
Man vermutet, daß die QSO ein kurzes *Entwicklungsstadium* der meisten Galaxien darstellen, das mit heftigen explosionsartigen Vorgängen im Kern verbunden ist. Das QSO-Stadium führt möglicherweise zum Stadium der Radiogalaxien.
Nach der Entdeckung der aktiven Prozesse in den Kernen vieler Galaxien wird die Möglichkeit erörtert, daß Galaxien nicht vor 10^{10} Jahren durch Kondensation von *Protogalaxien* aus der intergalaktischen Materie entstanden sind, sondern auch gegenwärtig noch aus den Kernen aktiver Galaxien als sehr kompakte Objekte mit Dichten über der Atomkerndichte ausgeworfen werden, wobei sie expandieren, normale Dichten erreichen und die kosmischen Erscheinungsformen der Materie bilden. Eine solche Theorie vertritt z. B. der sowjetische Astrophysiker AMBARZUMJAN.

10.4.3. Räumliche Verteilung

Längs des galaktischen Äquators in einer Breite von 20° bis zum Teil 40° werden keine Galaxien beobachtet. Das Fehlen von Galaxien in dieser *nebelfreien Zone* wird durch die lichtabsorbierende Wirkung der interstellaren Staubmaterie hervorgerufen. Nur in höheren galaktischen Breiten können wir die Sternsysteme mehr oder weniger ungeschwächt beobachten, weil ihr Licht diese Schicht nicht in ihrer vollen Ausdehnung durchlaufen muß.
Neben dieser vorgetäuschten gibt es auch reale Abweichungen von einer gleichmäßigen Verteilung der Galaxien im Raum. Man beobachtet Gruppen aus nur wenigen Mitgliedern und große Nebelhaufen, denen mehrere Hunderte Galaxien angehören.
10.4.3.1. Lokale Gruppe. Unsere Milchstraße bildet mit 15 benachbarten Galaxien eine physisch zusammengehörige Gruppe, die nur 3 Spiralsysteme, aber 9 elliptische Nebel enthält; mehr als die Hälfte der Mitglieder sind Zwergsysteme. Vermutlich gehören ihr noch mindestens 10 weitere Zwerggalaxien an, die erst in den letzten Jahren entdeckt wurden. Es ist nicht sicher, ob die Ansammlung über längere Zeit stabil ist. Der Durchmesser der Lokalen Gruppe beträgt rund 2 Mpc (Tab. 10.4.3-1).
10.4.3.2. Nebelhaufen. Im Mittel beobachtet man etwa 200 Mitglieder. Diese Zahl ist sicher zu klein, da nicht alle schwächeren Objekte erfaßt werden. Im Inneren eines Nebelhaufens enthält ein Raumwürfel von 1 Mpc etwa 500 bis

Tab. 10.4.3-2 Einige bekannte Nebelhaufen; N Anzahl der Galaxien, l und b galaktische Länge bzw. Breite in Grad, d° scheinbarer Durchmesser in Grad, r Entfernung in Mpc; V Radialgeschwindigkeit in km s^{-1}, P Dichte in Galaxien/Mpc^3, m_v scheinbare visuelle Helligkeit in Größenklassen

Haufen	N	l	b	d°	r	V	P	m_v
Virgo	2500	284	+74	12	11	+1150	500	9,4
Pegasus I	100	86	−48	1	40	3800	1100	12,5
Pisces	100	128	−29	10	40	5000	250	13,0
Cancer	150	202	+29	3	50	4800	500	13,4
Perseus	500	150	−14	4	58	5400	300	13,6
Coma	1000	80	+88	6	68	6700	40	13,5
Ursa Maj III	90	152	+64	0,7	80		200	14,5
Hercules		31	+44		105	10300		14,5
Pegasus II		84	−47			12800		15,2
Haufen A	400	144	−78	0,9	150	15800	200	16,0
Centaurus	300	313	+31	2	150		10	15,6
Ursa Maj I	300	140	+58	0,7	160	15400	100	16,0
Leo	300	232	+53	0,6	175	19500	200	16,3
Gemini	200	182	+19	0,5	175	23300	100	16,7
Cor. Bor.	400	41	+56	0,5	190	21600	250	16,3
Haufen B	300	345	−55	0,6	200		200	16,3
Bootes	150	50	+67	0,3	380	39400	100	18,0
Ursa Maj II	200	149	+54	0,2	380	41000	400	18,0
Hydra II		226	+30			60600		18,6

1000 Galaxien, d. h. 10000mal mehr als außerhalb. Die gegenseitigen Abstände der Galaxien sind im Mittel nur 6mal größer als ihre Durchmesser. Wegen ihrer relativen Geschwindigkeiten bis zu 10^6 ms^{-1} muß es häufig zu Zusammenstößen der Galaxien kommen, durch die die Sterne jedoch wegen ihrer relativ größeren Abstände nicht beeinflußt werden. Wohl aber treffen das interstellare Gas und der Staub unmittelbar aufeinander und verbleiben im gemeinsamen Schwerpunkt der Systeme, wenn sich die beiden Galaxien wieder getrennt haben. Da die Galaxien von der interstellaren Materie „leergefegt" wurden, kommt auch der Prozeß der Sternentstehung zum Stillstand. Die jungen, heißen, blauen Sterne verschwinden, die Farbe wird röter und das Spektrum „später". Da die Spiralstruktur offenbar eng mit der Anwesenheit interstellarer Materie verbunden ist, verschwinden auch die Spiralarme. Dies erklärt den Beobachtungsbefund, daß in den meisten Nebelhaufen fast nur elliptische Nebel und SO-Galaxien vorhanden sind (Tab. 10.4.3-2).
Einige Galaxienhaufen sind starke *Röntgenquellen*. Die Röntgenstrahlung ist thermische Bremsstrahlung des intergalaktischen Gases, das eine Temperatur von $2 \cdot 10^8$ K und eine Teilchendichte von $3 \cdot 10^{-6}$ cm^{-3} in den Haufen aufweist.

10.4.3.3. Großräumige Struktur. Man vermutet heute vielfach, daß die Gruppen- und Haufenbildung bei den Galaxien die Regel ist, daß Feldnebel nur Ausreißer sind, die die Gravitationskraft des Haufens wegen ihrer hohen Geschwindigkeit überwinden konnten. Es wird weiter vermutet, daß noch weit großräumigere Gesamtballungen, die *Superhaufen*, existieren, die mehrere Nebelhaufen, Gruppen und Feldgalaxien umfassen und ein riesiges physikalisches System bilden. Auch die Lokale Gruppe könnte Mitglied des Lokalen Superhaufens sein, dessen Zentrum im großen Virgohaufen liegen müßte. Die Umlaufszeit um dieses Zentrum würde etwa 50 bis 200 Mill. Jahre betragen.
Eng verknüpft mit diesen Vermutungen ist die für die Kosmologie grundlegende Frage, ob die Materie im Weltall trotz aller Gruppen- und Haufenbildungen homogen und isotrop, d. h. in allen Richtungen und Entfernungen gleichmäßig verteilt ist; ob es möglich ist, wahllos in allen Richtungen gleichgroße Raumteile herauszugreifen, in denen die Anzahl der Galaxien stets gleich ist. Offenbar müßten diese Raumgebiete so groß sein, daß die durch Haufen oder mögliche Superhaufen gebildeten Schwankungen in der Nebeldichte keine Rolle mehr spielen. Gestützt auf Beobachtungen an der 3-K-Strahlung (↑ 10.4.4.1.) ist man der Ansicht, daß das Weltall gegenwärtig in diesem Sinne homogen und isotrop aufgebaut ist. Diese Annahme wird *kosmologisches Prinzip* oder *Postulat* genannt.
Danach müßte das Weltall von allen Punkten den gleichen Anblick bieten, wenn man von den örtlichen Ansammlungen in Form von Galaxiengruppen und Haufen absieht. Das kosmologische Prinzip bedeutet eine Verallgemeinerung der Kopernikanischen Idee, daß der Kosmos keinen Mittelpunkt oder ausgezeichneten Punkt hat.
Eine Möglichkeit die Verteilung der Objekte im Kosmos festzustellen, bietet die statistische Methode der *Nebelzählung*. Eine Kugel mit dem Radius r enthält bei gleichmäßiger Verteilung $N \sim r^3$ Objekte. Jedes Objekt erzeugt im Mittelpunkt der Kugel eine Strahlungsintensität von $S \sim 1/r^2$. Man „sieht" deshalb im Zentrum bis zur Helligkeit S insgesamt $N \sim S^{-3/2}$ Objekte.

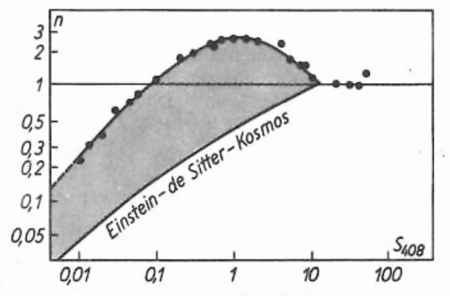


Abb. 10.4.3-3 Anzahl der Radioquellen bei verschiedenen Flußdichten (Erläuterung im Text)

Dieses *N-S-Gesetz* wird heute vorwiegend auf Radioquellen angewandt, da Radioteleskope eine höhere Reichweite als optische Teleskope besitzen. Man beobachtet, daß die Zahl der Radioquellen, mit schwächer werdendem S, d. h. mit wachsender Entfernung, stärker zunimmt als sich aus einer gleichmäßigen Verteilung ergibt. Erst bei sehr schwachen Flußdichten S nimmt die Zahl der Objekte wieder ab. Abb. 10.4.3-3 zeigt die Anzahl der Radioquellen bei verschiedenen Flußdichten S_{408} in 10^{-26} Wm^{-2}Hz^{-1} nach Messungen des Cambridger Mullard Observatoriums bei einer Frequenz von 408 MHz. Die Anzahl n ist auf den hypothetischen Verlauf $N \sim S^{-3/2}$ (↑ 10.4.3.3.) in einem statischen euklidischen Kosmos mit konstant vorausgesetzter räumlicher Dichte der Radioquellen bezogen (horizontale Linie bei $n = 1$). Bei kleinen Radioflußdichten erfaßt man entfernte Quellen; bei $S_{408} \geqq 10$ liegen vorwiegend benachbarte Quellen. In expandierenden Weltmodellen (↑ 10.4.5.1.) muß n auch bei konstanter räumlicher Dichte wegen der Lichtschwächung durch die Rotverschiebung rasch mit abnehmendem S fallen. Als Beispiel ist der theoretische Verlauf im *Einstein-de-Sitter-Kosmos* eingezeichnet. Der beobachtete Überschuß (grau) ist nur durch eine große Häufung von

Radiogalaxien und Quasaren in der Vergangenheit des Kosmos zu erklären.
Wegen der Laufzeit des Lichts bedeuten wachsende Entfernungen einen immer weiter reichenden Blick in die Vergangenheit des Kosmos. Man muß daher diesen Beobachtungsbefund dahingehend deuten, daß die Zahl der Radiogalaxien und Quasare in der Vergangenheit weitaus größer war als heute. Bis zu einer Rotverschiebung von $z \approx 4$ nimmt die Anzahl der extragalaktischen Radioquellen in der Raumeinheit proportional $(1 + z)^6$ zu.
Für das überschaubare Weltall, gelegentlich als *Metagalaxis* bezeichnet, werden heute folgende Daten angenommen: Die Gesamtzahl der Galaxien bis zu einer Entfernung von 3000 Mpc oder 10 Mrd. ly wird auf 3 Mrd. geschätzt, das bedeutet etwa 1 Galaxie von der Größe der Milchstraße in einem Würfel von 3 Mpc Kantenlänge und eine mittlere Dichte der Materie von 10^{-31} bis 10^{-30} gcm^{-3}. Die Gesamtmasse in diesem Bereich dürfte bei 10^{50} kg liegen.

10.4.4. Thermodynamik des Universums

Die aus der Rotverschiebung und der Hubblebeziehung (↑ 10.4.2.1.) sich ergebende Expansion des Weltalls bedeutet, daß der Kosmos nicht zu allen Zeiten den gleichen Anblick bieten kann. So muß z. B. die mittlere Massendichte mit der Zeit abnehmen. Andererseits muß sich auch das Verhältnis der Anteile von stofflicher Materie und Strahlung mit der Zeit verändern.

10.4.4.1. 3-Kelvin-Strahlung. Im Jahre 1965 entdeckten PENZIAS und WILSON im Zentimeterwellenbereich radioastronomisch eine isotrop über den ganzen Himmel verteilte Strahlung, die man nicht den Galaxien zuordnen konnte. Es handelt sich dabei um eine reine Temperaturstrahlung, deren Intensitätsmaximum im Millimeterlängenbereich liegt und der Temperatur von 3 K entspricht (Abb. 10.4.4-1). Die Strahlungsdichte beträgt 10^3 Photonen pro cm^3, d. h., auf ein Teilchen der stofflichen Materie (Baryonen und Leptonen) kommen rund 10^{10} Photonen. Allerdings ist die Energie der Photonen so gering, daß die Energiedichte der 3-Kelvin-Strahlung gegenwärtig nur rund 1/1000 der Ruhmassenenergiedichte im Kosmos beträgt. Diese kosmische Hintergrundstrahlung ist ein „Relikt" aus der Vergangenheit des Kosmos.

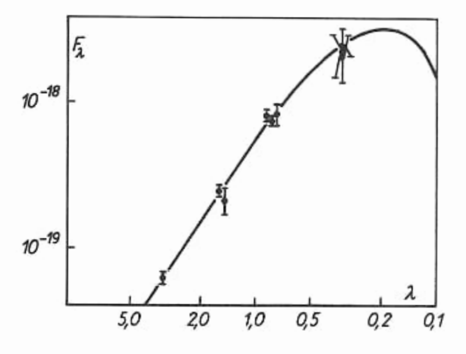


Abb. 10.4.4-1 Spektrum der 3-Kelvin-Strahlung; die beobachtete Radioflußdichte F_λ in Wm^{-2} sr^{-1} Hz^{-1} bei verschiedenen Wellenlängen λ in cm läßt sich durch eine Plancksche Temperaturstrahlung von 3 K darstellen (*ausgezogene Linie*)

10.4.4.2. Intergalaktische Materie. Neben leuchtenden Brücken und Filamenten, die sich weit über die Galaxien in den Raum hinaus erstrekken, und dem durch die Röntgenstrahlung in Galaxienhaufen nachgewiesenen intergalaktischen, heißen Gas gibt es keine sicheren Hinweise für das Vorhandensein eines intergalaktischen Mediums. Zumindest ist dessen Dichte gegenwärtig vernachlässigbar gering gegenüber der in Galaxien konzentrierten Masse.
Die *Röntgenhintergrundstrahlung* läßt jedoch die Schlußfolgerung zu, daß in großen Entfernungen, d. h. in der Vergangenheit des Kosmos, noch eine beträchtliche Menge heißes intergalaktisches Gas (vorwiegend Wasserstoff) vorhanden gewesen sein muß. Dieses Gas wurde dann entweder bei der Galaxienbildung aufgebraucht oder von den Galaxienhaufen eingefangen.

10.4.4.3. Evolution des Kosmos. Gegenwärtig sind die Hauptenergiebeiträge der Materie im Kosmos, die Strahlungsenergie und die Ruhmassenenergie, entkoppelt. Bei der Expansion ändern sich die Energiedichten beider Bestandteile unterschiedlich, so daß sich ihr Verhältnis verändert. Bezeichnet $R(t)$ als zeitabhängiger Skalenfaktor die durch die Expansion bedingte Veränderung der gegenseitigen Abstände im Kosmos (↑ 10.4.5.1.), dann verlangt der Erhaltungssatz der Baryonen und Leptonen für die Ruhmassenenergiedichte $\varrho_M \sim R^{-3}$; für die Strahlungstemperatur T_s im Prozeß der adiabatischen Abkühlung gilt $T_s \sim R^{-1}$ bzw. für die Strahlungsenergiedichte $u_s \sim R^{-4}$. Um einheitliche Maßeinheiten zu haben, wird häufig u_s in Jcm^{-3} formal gemäß der Beziehung $u_s/c^2 = \varrho_s$ in ein Masseäquivalent, z. B. in gcm^{-3}, umgerechnet (Abb. 10.4.4-2). Diese Beziehungen erlauben es nun, das Verhältnis der Energieanteile in die Vergangenheit des Kosmos zu verfolgen.
Abb. 10.4.4-2 verdeutlicht, daß die *Energiedichte der Strahlung* ϱ_s rascher abnimmt, als die *Ruhmassenenergiedichte* ϱ_M der stofflichen Materie. Für beide Größen gilt die rechte Ordinate.

ϱ_c ist die *kritische Dichte* (↑ 10.4.5.1.). Die Temperatur T_s der kosmischen Hintergrundstrahlung, vor der Entkopplung auch die des kosmischen Mediums, zeigt die linke Ordinate. In den ersten 1000 s nach der Singularität liefen zahlreiche Kernreaktionen ab; in dieser Epoche wurde auch die beobachtete kosmische Häufigkeit des Heliums erzeugt. Objekte, deren Licht so lange unterwegs war, daß wir sie in einer

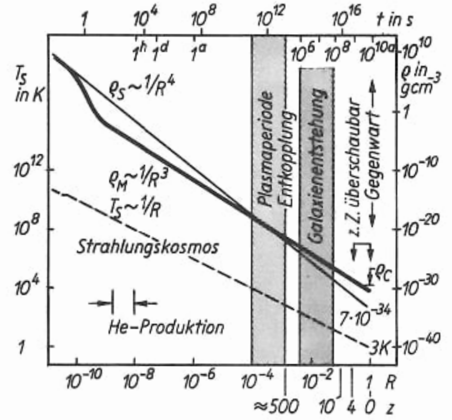


Abb. 10.4.4-2 Geschichte des Kosmos; zeitliche Entwicklung *oben* in Sekunden bis Jahren, *unten* nach dem Wert des Skalenfaktors (Erläuterung im Text)

durch t oder R gekennzeichneten vergangenen Epoche sehen, hätten die unten angegebene *Rotverschiebung* z. Die logarithmische Darstellung überbetont die Frühphasen des Kosmos. Von der seit der Singularität bis heute verstrichenen Zeit von 12 bis 15 Mrd. Jahren umfaßt nur 1/100 die Spanne bis zum Ende der Epoche der Galaxienentstehung.

Geht man von der Gegenwart in die Vergangenheit, so gelangt man bei einer Temperatur von 2000 bis 4000 K in die *Plasmaperiode*. Der interstellare und intergalaktische Wasserstoff war ionisiert; zwischen Gas und Strahlung bestand thermodynamisches Gleichgewicht, weil thermische Energie und Strahlungsenergie stark gekoppelt waren. Zu diesem Zeitpunkt waren auch Ruhmassenenergiedichte und Strahlungsenergiedichte gleich. Vorher überwog die Strahlungsenergiedichte. Oberhalb einer Temperatur von 10^{10} K war die Energie der Photonen so groß, daß sie zur Paarerzeugung (Elektron + Positron) führte. Damit erreichte die Ruhmassenenergiedichte der stofflichen Materie wieder die Strahlungsenergiedichte. Über 10^{12} K gab es ein Gleichgewicht zwischen allen Energiearten. In den ersten Sekunden nach der kosmologischen Singularität (↑ 10.4.5.1.) kam es zu zahlreichen Kernreaktionen. Geht man davon aus, daß ursprünglich nur Protonen und Neutronen anwesend waren, so ergeben die Rechnungen für diese Phase, daß nur Helium in merklicher Menge damals gebildet worden war. Sie erklären die heute beobachtete Häufigkeit von 25 bis 30 % Masseanteil, die nicht durch die Wasserstoffverbrennung in Sternen entstanden sein kann (↑ 10.3.1.).

Die Phase der *Galaxienentstehung* setzte vermutlich 10^6 bis 10^8 Jahre nach der Singularität ein. Zwei Mechanismen werden gegenwärtig erörtert: das Anwachsen von Störungen in einem sonst homogenen Medium bis zu gravitativen Instabilitäten und nachfolgendem Kollaps, und der Auswurf aus bereits vorhandenen hochverdichteten Objekten. Eine befriedigende Lösung dieser Problematik ist bisher noch nicht gefunden worden.

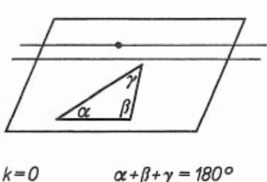
10.4.5. Kosmologie

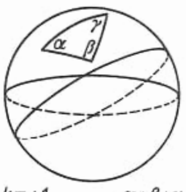
Die Kosmologie untersucht den Zustand des Kosmos als Ganzes, seinen Aufbau und seine Entwicklung. Eine kosmologische Theorie muß davon ausgehen, daß die bekannten physikalischen Gesetze im Weltall universell, d. h. räumlich und zeitlich unverändert gelten. Wohl wird es bisher noch unbekannte Gesetze geben; Spekulationen darüber dürfen aber nicht zu Widersprüchen mit fundierten physikalischen Erkenntnissen führen. Man versuchte deshalb, mit möglichst wenig einschneidenden Annahmen aus den physikalischen Grundgesetzen und aus Hypothesen *Weltmodelle*, die den heutigen Zustand der Welt beschreiben, zu konstruieren und durch Vergleich mit den Beobachtungsdaten zu entscheiden, ob sie der Wirklichkeit entsprechen. Meist wird stillschweigend vorausgesetzt, daß die Materie im Kosmos kontinuierlich verteilt ist. Die Betrachtung großer Dimensionen rechtfertigt diese Annahme. Analog wird ein Gas makroskopisch auch als homogenes Medium behandelt und dessen atomare oder molekulare Struktur übersehen. Unterscheidungsmerkmale zwischen „vernünftigen" Weltmodellen werden im allgemeinen erst in großen Entfernungen, bei den schwächsten Galaxien, merkbar. Von ihnen ist das Licht Milliarden Jahre unterwegs; man sieht diese Objekte in einem wesentlich früheren Zustand ihrer Entwicklung und weiß daher nicht, ob ihre integralen Eigenschaften, wie Leuchtkraft und Farbe, mit denen benachbarter Galaxien vergleichbar sind.

Bei der Behandlung kosmologischer Fragen dringt man in Raumdimensionen vor, in denen die alltäglichen Vorstellungen versagen. Daß insbesondere unsere Raumanschauung der täglichen Erfahrung nicht ohne weiteres auf die unendlichen Dimensionen des Kosmos über-

tragen werden kann, läßt sich an zwei Beispielen zeigen:
Angenommen der Raum sei gleichmäßig mit Sternen oder Galaxien erfüllt, die gleiche Leuchtkraft haben sollen. In der Entfernung r haben sie dann die scheinbare Helligkeit $I(r) \sim 1/r^2$, ihre Anzahl ist aber $N(r) \sim r^3$. Aus einem Himmelsareal erhält man daher von den Sternen in der Entfernung r die Strahlungsleistung $I(r) \cdot N(r) \sim r$. Summiert man über alle Entfernungen, so wächst dieser Ausdruck über alle Grenzen für $r \to \infty$ und der Himmel müßte nachts strahlend hell erscheinen. Man spricht vom *Olbersschen Paradoxon.*
Auch die Gravitationskraft eines Sterns am Ort der Erde ist proportional $1/r^2$, und die Anzahl der Sterne wächst mit r^3, so daß die Summe der Kräfte aller Sterne aus einem Himmelsareal unendlich groß ist. Das Gravitationspotential an jedem Ort im Weltall wäre nach diesem *Seeligerschen Paradoxon* völlig unbestimmt.
Ohne auf die Problematik der Beispiele näher einzugehen, läßt sich doch sagen, daß die gezeigten Schwierigkeiten sowohl durch einen *hierarchischen Aufbau* der Welt, bei dem jedes System nur Teil eines größeren übergeordneten Systems ist, das wiederum nur Untersystem der nächsthöheren Ordnung ist, als auch durch die *Rotverschiebung* beseitigt werden können.
Die moderne Kosmologie beruht auf der *allgemeinen Relativitätstheorie* von Einstein und wird als *relativistische Kosmologie* bezeichnet. Sie liefert eine konsequente und widerspruchsfreie Beschreibung der Raum-Zeit-Struktur des Weltalls. Raum und Zeit sind Existenzformen der Materie, und die im Weltall verteilte Materie tritt als bestimmende Größe für die Geometrie des Raums und seine zeitliche Entwicklung auf.

10.4.5.1. Relativistische Kosmologie. Relativistische Weltmodelle unterscheiden sich durch die Eigenschaften des Raums. Unterschiede zwischen Räumen verschiedener Krümmung k lassen sich im zweidimensionalen Fall, für Flächen, leicht kennzeichnen. In dem *euklidischen Raum* ist eine Fläche mit $k = 0$ eine Ebene. In ihr gibt es zu einer Geraden durch einen nicht auf dieser Geraden liegenden Punkt nur eine Parallele; die Winkelsumme in einem Dreieck beträgt immer 180°, und der Flächeninhalt einer Kreisfläche wächst mit dem Radius unbegrenzt.
Der Raum der *elliptischen* oder *Riemannschen Geometrie* hat positive Krümmung. Als Veranschaulichung einer Fläche mit $k = \text{const.} > 0$ kann die Kugeloberfläche dienen; ihre Großkreise entsprechen den Geraden. Zwei dieser Geraden schneiden sich stets, es gibt keine Parallelen. Die Winkelsumme in einem solchen Dreieck ist größer als 180°. Der Flächeninhalt von beliebigen Kreisen bleibt stets endlich (Abb. 10.4.5-1).
Der Raum der *hyperbolischen* oder *Lobatschewskischen Geometrie* hat eine negative Krümmung. Das zweidimensionale Analogon einer Fläche mit $k = \text{const.} < 0$ ist die *Pseudosphäre*, eine Fläche, die durch Rotation einer Traktrix um ihre Asymptote entsteht. Zu einer Geraden gibt es dort unendlich viele Parallelen, die sich nirgends im Endlichen schneiden; die Winkelsumme in diesem Dreieck ist stets kleiner als 180°, und wie in der euklidischen Geometrie gibt es auf dieser Fläche keinen größten endlichen Flächeninhalt.
In kleineren Bereichen können Flächenstücke als eben angesehen werden; in ihnen gilt die euklidische Geometrie. Deshalb können im täglichen Leben, wie auch in etwas größeren Gebieten des Weltalls, die Lehrsätze der Schulgeometrie, beispielsweise der Satz des Pythagoras, ohne merkliche Fehler angewandt werden.
Als Abstand ds zweier benachbarter Punkte erhält man in der Ebene $ds^2 = dx^2 + dy^2$ oder in Polarkoordinaten $ds^2 = dr^2 + r^2(d\Theta)^2$; auf einer Kugeloberfläche mit dem Radius a dagegen $ds^2 = a^2(dr^2 + r^2\, d\Phi^2)/(1 + r^2/4)^2$ und auf einer hyperbolischen Fläche ergibt sich $ds^2 = a^2(dr^2 + r^2\, d\Phi)^2/(1 - r^2/4)^2$. Man nennt diese Ausdrücke koordinatendifferentiale *Metrik*, da sie aus der Maßbestimmung des Raums folgen. Für drei Dimensionen erhält man $ds^2 = g_1\, dx^2 + g_2\, dy^2 + g_3\, dz^2$ und nennt die Funktionen g_i Komponenten des *metrischen Tensors*. Natürlich hängt ihre äußere Form von der speziellen Wahl des Koordinatensystems ab; es läßt sich aber nachweisen, daß zwischen den drei oben angegebe-





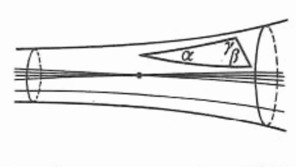


Abb. 10.4.5-1 Fläche mit euklidischer ($k = 0$), elliptischer ($k = +1$) und hyperbolischer Geometrie ($k = -1$)

nen Maßbestimmungen so grundsätzliche Unterschiede bestehen, daß sie sich durch keine Koordinatentransformationen ineinander überführen lassen.

In der relativistischen Kosmologie kommt zu den drei Raumkoordianten noch die Zeit t als vierte Koordinate hinzu. Als „Abstand" ds gilt das Intervall zwischen zwei „Ereignissen" mit den Koordinaten (t_1, x_1, y_1, z_1) und (t_2, x_2, y_2, z_2). Das kosmologische Postulat führt dann in mathematischer Formulierung auf die *Robertson-Walker-Metrik*, in Polarkoordinaten: $\mathrm{d}s^2 = c^2\,\mathrm{d}t^2 - R^2(t)\cdot[\mathrm{d}r^2 + r^2(\mathrm{d}\Theta^2 + \sin^2\Theta\,\mathrm{d}\Phi^2)]/(1 + kr^2/4)^2$. Die Konstante k kennzeichnet die *Krümmung des Raums* und ist für den elliptischen Raum $+1$, für den euklidischen 0 und für den hyperbolischen -1. Der *Skalenfaktor* $R(t)$ ist nur eine Funktion der Zeit und beschreibt die zeitliche Veränderung der räumlichen Abstände. In einem statischen Universum wäre R eine Konstante; c bedeutet die Lichtgeschwindigkeit; r ist eine zeitlich unveränderliche, radiale, mitbewegte Koordinate. Die physikalische Entfernung errechnet sich aus r *und* $R(t)$. Wenn z die beobachtete Rotverschiebung eines extragalaktischen Objekts ist, dessen Licht in der Vergangenheit zu einem Zeitpunkt t_1 ausgesandt wurde, als der Skalenfaktor den Wert $R(t_1) = R_1$ hatte und R_0 seinen gegenwärtigen Wert bezeichnet, so besteht zwischen z und $R(t)$ die einfache Beziehung $1 + z = R_0/R_1$. R_0 kann gleich Eins gesetzt werden, so wie das in Abb. 10.4.4-2 der Fall ist.

Es war die große Leistung von EINSTEIN, diese Geometrie des Raums mit den Eigenschaften der darin befindlichen Materie zu verknüpfen. Aus seinen Feldgleichungen ergeben sich die fundamentalen Differentialgleichungen für das kosmologische P oblem (ohne die kosmologische Konstante): $8\pi G\varrho = (3/R^2)\,[k + (\mathrm{d}R/\mathrm{d}t)^2]$ und $8\pi Gp/c^2 = -(2/R)\,(\mathrm{d}^2R/\mathrm{d}t^2) - (1/R^2)\,(\mathrm{d}R/\mathrm{d}t)^2 - k/R^2$. Hierbei bedeuten G die Gravitationskonstante, ϱ die räumliche Dichte der Materie und p ihren Druck. Auf der rechten Seite der Gleichungen stehen der Skalenfaktor und seine Ableitungen nach der Zeit sowie die Krümmungskonstante k. Es ist nun die Aufgabe der Kosmologie, für den oeobachteten Kosmos eine passende Lösung für diese Gleichungen zu finden. Die Lösungen geben uns Auskunft über die Geometrie des Weltalls und über die zeitliche Entwicklung, d. h. über den Skalenfaktor $R(t)$. Der Druck p wird bei diesen Lösungen im allgemeinen vernachlässigt, denn er beinhaltet im wesentlichen den Strahlungsdruck. Im gegenwärtigen Zustand des Weltalls ist p/c^2 gegenüber der Dichte sehr klein.

Jede Lösung der Gleichungen ist ein *Weltmodell*. Weltmodelle mit positiver Krümmung sind räumlich geschlossen; sie haben einen endlichen Rauminhalt, sind aber wie eine Kugeloberfläche im zweidimensionalen Fall unbegrenzt, wenn man in einer Richtung fortschreitet. Räumlich unendliche Modelle gibt es für den euklidischen Fall, den Einstein-de-Sitter-Kosmos $k = 0$, und den hyperbolischen Fall. Die zeitliche Entwicklung dieser Modelle erhält man aus der Funktion $R(t)$. Alle Modelle zeigen eine Expansion, die von einem singulären Punkt unendlich großer

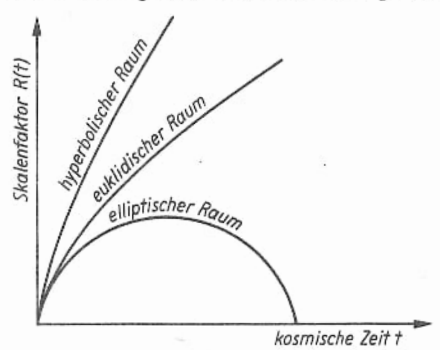


Abb. 10.4.5-2 Zeitlicher Verlauf des Skalenfaktors für drei Weltmodelle

Dichte und verschwindender Ausdehnung ($R \approx 0$), der *kosmologischen Singularität*, ausgeht und entweder bis ins Unendliche in unendlich langer Zeit oder bis zu einer maximalen Ausdehnung führt, von der aus eine Kontraktion wieder zu einer Singularität in endlicher Zeit führt.

Der sogenannte „Urknall" liegt nur dann in unendlich ferner Vergangenheit, wenn man in die Einsteinschen Feldgleichungen neben der Newtonschen Anziehungskraft eine Abstoßungskraft, das *kosmologische Glied*, einführt. Da die Expansion verzögert erfolgt (Abb. 10.4.5-2), ist die seit der kosmologischen Singularität verstrichene Zeit, das *Friedmannalter*, in allen Modellen kleiner als die reziproke Hubblekonstante $1/H = 18$ Mrd. Jahre, das *Hubblealter*. Dieses verhältnismäßig kurze Weltalter kennzeichnet das Entwicklungsalter des beobachtbaren Kosmos. Ein vergleichbares Alter haben die zur Population II gehörenden Objekte in unserer Galaxis und in anderen Galaxien. Wesentlich ältere Objekte kennen wir nicht. In der Nähe der Singularität sind die Voraussetzungen zur Lösung der kosmologischen Gleichungen nicht mehr gültig. Hier werden sicher Quanteneffekte der Gravitation eine Rolle spielen, für die eine Theorie noch aussteht, so daß Aussagen über die Singularität selbst sowie über die Zeit vor der Singularität nicht möglich sind.

Um nachzuprüfen, welches Modell die Wirklichkeit widerspiegelt, werden aus den kosmologischen Differentialgleichungen Beziehungen zwischen der Geometrie und beobachtbaren

Größen abgeleitet. So ist z. B. im Fall des euklidischen Einstein-de-Sitter-Kosmos die mittlere Massendichte mit der Hubblekonstante $H = (1/R)(\mathrm{d}R/\mathrm{d}t)$ gemäß $\varrho_c = (3H^2/8\pi G)$ verknüpft. Der Wert der *kritischen Dichte* ϱ_c beträgt $5 \cdot 10^{-30}$ gcm^{-3}. Die Hubblekonstante (↑ 10.4.2.1.) verändert sich mit der Zeit; ihr gegenwärtiger Wert 50 kms^{-1}Mpc^{-1} sei mit H_0 bezeichnet. Krümmungskonstante k, Hubblekonstante H und mittlere Massendichte ϱ sind gemäß der Beziehung $k = H^2R^2(\varrho/\varrho_c - 1)$ miteinander verknüpft. Je nachdem, ob die räumliche Dichte größer oder kleiner als die kritische Dichte ist, ist der Raum elliptisch oder hyperbolisch gekrümmt. Die Massenbestimmungen im Kosmos ergaben für die mittlere Massendichte ϱ einen Wert, der unterhalb der kritischen Dichte liegt, so daß der hyperbolische Fall verwirklicht zu sein scheint.

Eine weitere wichtige Bestimmungsgröße, die aber nicht unabhängig vom Dichteparameter ϱ/ϱ_c ist, ist die Beschleunigungszhal $q = -[1/(RH^2)]\,[\mathrm{d}^2R/\mathrm{d}t^2]$, eine Kombination aus der zweiten Ableitung des Skalenfaktors $R(t)$ nach der Zeit und der Hubblekonstanten. Diese Zahl ist nur in speziellen Modellen eine Konstante und nicht mehr zeitabhängig; ihr gegenwärtiger Wert sei q_0. Aus den Differentialgleichungen kann man dann sehr leicht folgende Beziehungen zwischen der Geometrie und den oben genannten Parametern für ein Weltall mit $p = 0$ ableiten:

$\varrho_0 = [3H_0^2/(4\pi G)]\,q_0, \; k/R_0^2 = H_0^2(2q_0 - 1)$.

Nach der zuletzt angeführten Gleichung muß wegen $k = +1$ für geschlossene elliptische Modelle $q_0 > 1/2$ sein, wegen $k = 0$ für einen euklidischen Raum $q_0 = {}^1/_2$ und wegen $k = -1$ für Modelle mit hyperbolischer Geometrie $0 \leqq q_0 < {}^1/_2$. Für $q_0 = 0$ wäre nach der obigen ersten Gleichung $\varrho_0 = 0$, das Weltall demzufolge leer.

Die *Beschleunigungszahl q* und damit die Geometrie des Raums läßt sich im Prinzip auch aus der Abhängigkeit der Rotverschiebung von der scheinbaren Helligkeit bestimmen (↑ Abb. 10.4.2-2). Blickt man nämlich weit in den Raum hinaus und damit weit in die Vergangenheit zurück, so müßte wegen der Verzögerung der Expansion die Geschwindigkeit der Galaxien größer sein als sich aus der Proportionalität zwischen Entfernung und gegenwärtigem Wert der Hubblekonstanten ergibt. Man hat allerdings dabei zu berücksichtigen, daß sich wegen der Rotverschiebung auch die scheinbaren Helligkeiten merklich ändern, weil unterschiedliche Teile des Spektrums in den optischen Bereich gelangen. Die Beziehung zwischen Rotverschiebung und Helligkeit fächert sich für verschiedene q_0-Werte in eine Kurvenschar auf, die allerdings erst unterhalb der 19. Größe für normale Galaxien merkbare Unterschiede ergeben. Gegenwärtig sind die Beobachtungen noch zu ungenau, so daß nach dieser Methode eine Entscheidung über die Geometrie des Raums getroffen werden kann.

Eine bemerkenswerte Eigenschaft vieler kosmologischer Modelle ist die Existenz eines *Horizonts* in einer endlichen Entfernung, über den man prinzipiell nicht hinausblicken kann. Galaxien am Horizont haben unendlich große Rotverschiebung, sind also nicht mehr sichtbar. Der Horizont rückt aber mit der Zeit immer weiter weg, so daß nach und nach immer entferntere Teile des Weltalls zugänglich werden. Eine weitere bemerkenswerte Eigenschaft vieler Modelle ist, daß die Winkelgröße der Galaxien zunächst mit der Entfernung abnimmt, jedoch vor dem Horizont ein Minimum erreicht und dann wieder anwächst.

Wegen der Verflechtung von Expansion und Lichtausbreitung sind Entfernungsangaben abhängig vom zugrunde gelegten Modell sowie von der Methode, nach der die Entfernung bestimmt werden soll. Häufig gibt man

$R_0 \int \frac{\mathrm{d}r}{1 + kr^2/4}$ als den nichtmeßbaren *Lichtweg* an, wo r die radiale Polarkoordinate (mitbewegte Koordinate) aus der Robertson-Walker-Metrik bedeutet, die sich zeitlich nicht verändert.

Der Lichtweg bedeutet die Strecke, die das Licht im gegenwärtigen Zustand des Weltalls vom Objekt bis zu uns zurückzulegen hätte. Um diese Strecke zu realisieren, müßte man die Expansion anhalten. Andere Entfernungsangaben beziehen sich auf die scheinbare Helligkeit, d. h. die *fotometrische Entfernung*, oder auf die scheinbare Winkelgröße der Galaxien. Alle Entfernungen hängen vom Modell ab und sind nur für Entfernungen bis 1 Gpc einander gleich.

Erweitert man das kosmologische Postulat dahingehend, daß der Kosmos auch zu *allen Zeiten* den gleichen Anblick bietet, so gelangt man zur Theorie des *stationären Weltalls*, der *Steady-State-Theorie*. Da die räumliche Dichte nach dieser Theorie zeitlich konstant bleiben muß, wäre wegen der wachsenden Verdünnung durch die Expansion eine ständige Erzeugung von Materie aus dem Nichts nötig. Gefordert wird die Entstehung von einem einzigen Wasserstoffatom im Jahr in einem Raumgebiet von 10 km^3, um die Voraussetzungen zu erfüllen. Da diese geringe Menge nicht nachweisbar ist, bleiben die Erhaltungssätze von Masse und Energie als Erfahrungssätze bestehen. Die Beobachtungen lassen diese Theorie allerdings wenig wahrscheinlich erscheinen. So war die räumliche Dichte der extragalaktischen Radioquellen in der Vergangenheit größer als heute (↑ 10.4.3.3.). Auch die 3-Kelvin-Strahlung (↑ 10.4.4.1.) ist

schwer mit der Steady-State-Theorie zu vereinbaren.

10.4.5.2. Ausblick. Sicherlich werden die vorstehend umrissenen Prinzipien des Aufbaus und der Entwicklung der nichtbelebten Materie mit dem ständig fortschreitenden Erkenntniszuwachs noch manche Modifikation erfahren. Gegenwärtig scheint es, daß eine endliche, wenn auch sehr kleine Ruhmasse des Neutrinos (↑ 10.3.3.7.) die Dichterelationen der kosmischen Materie wesentlich verschieben würde. Man vermutet neuerdings, daß ein erheblicher, ja sogar überwiegender Anteil der kosmischen Materie von Neutrinos beigetragen wird. Dadurch ließe sich erklären, warum Galaxiensysteme mit relativ „dünner" Konzentration an konventioneller Materie dennoch durch die Gravitation zusammengehalten werden.

Möglicherweise werden bei Erscheinen dieses Buches hierzu schon neue Ergebnisse von Beobachtungen und theoretischen Modellbetrachtungen vorliegen.

Ordnungs-zahl	Element	Symbol	Elektronen-konfiguration	Ionisations-energie in eV
1	Wasserstoff	H	$1s$	13,53
2	Helium	He	$1s^2$	24,56
3	Lithium	Li	He + $2s$	5,37
4	Beryllium	Be	He + $2s^2$	9,48
5	Bor	B	He + $2s^2 2p$	8,4
6	Kohlenstoff	C	He + $2s^2 2p^2$	11,22
7	Stickstoff	N	He + $2s^2 2p^3$	14,47
8	Sauerstoff	O	He + $2s^2 2p^4$	13,56
9	Fluor	F	He + $2s^2 2p^5$	18,6
10	Neon	Ne	He + $2s^2 2p^6$	21,5
11	Natrium	Na	Ne + $3s$	5,12
12	Magnesium	Mg	Ne + $3s^2$	7,61
13	Aluminium	Al	Ne + $3s^2 3p$	5,96
14	Silicium	Si	Ne + $3s^2 3p^2$	7,39
15	Phosphor	P	Ne + $3s^2 3p^3$	10,3
16	Schwefel	S	Ne + $3s^2 3p^4$	10,31
17	Chlor	Cl	Ne + $3s^2 3p^5$	12,96
18	Argon	Ar	Ne + $3s^2 3p^6$	15,69
19	Kalium	K	Ar + $4s$	4,32
20	Calcium	Ca	Ar + $4s^2$	6,09
21	Scandium	Sc	Ar + $4s^2 3d$	6,57
22	Titanium	Ti	Ar + $4s^2 3d^2$	6,80
23	Vanadium	V	Ar + $4s^2 3d^3$	6,76
24	Chromium	Cr	Ar + $4s 3d^5$	6,74
25	Mangan	Mn	Ar + $4s^2 3d^5$	7,40
26	Eisen	Fe	Ar + $4s^2 3d^6$	7,83
27	Cobalt	Co	Ar + $4s^2 3d^7$	7,81
28	Nickel	Ni	Ar + $4s^2 3d^8$	7,61
29	Kupfer	Cu	Ar + $4s 3d^{10}$	7,69
30	Zink	Zn	Ar + $4s^2 3d^{10}$	9,35
31	Gallium	Ga	Ar + $4s^2 3d^{10} 4p$	5,97
32	Germanium	Ge	Ar + $4s^2 3d^{10} 4p^2$	7,85
33	Arsen	As	Ar + $4s^2 3d^{10} 4p^3$	9,4
34	Selen	Se	Ar + $4s^2 3d^{10} 4p^4$	9,75
35	Brom	Br	Ar + $4s^2 3d^{10} 4p^5$	11,80
36	Krypton	Kr	Ar + $4s^2 3d^{10} 4p^6$	13,94
37	Rubidium	Rb	Kr + $5s$	4,16
38	Strontium	Sr	Kr + $5s^2$	5,67
39	Yttrium	Y	Kr + $5s^2 4d$	6,5
40	Zirconium	Zr	Kr + $5s^2 4d^2$	6,95
41	Niobium	Nb	Kr + $5s 4d^4$	6,77
42	Molybdän	Mo	Kr + $5s 4d^5$	7,35
43	Technetium	Tc	Kr + $5s 4d^6$	7,45
44	Ruthenium	Ru	Kr + $5s 4d^7$	7,7
45	Rhodium	Rh	Kr + $5s 4d^8$	7,7
46	Palladium	Pd	Kr + $4d^{10}$	8,5
47	Silber	Ag	Kr + $5s 4d^{10}$	7,54
48	Cadmium	Cd	Kr + $5s^2 4d^{10}$	8,95
49	Indium	In	Kr + $5s^2 4d^{10} 5p$	5,76
50	Zinn	Sn	Kr + $5s^2 4d^{10} 5p^2$	7,37
51	Antimon	Sb	Kr + $5s^2 4d^{10} 5p^3$	8,5
52	Tellur	Te	Kr + $5s^2 4d^{10} 5p^4$	8,89
53	Iod	I	Kr + $5s^2 4d^{10} 5p^5$	10,44
54	Xenon	Xe	Kr + $5s^2 4d^{10} 5p^6$	12,08

Ordnungs-zahl	Element	Symbol	Elektronen-konfiguration	Ionisations-energie in eV
55	Caesium	Cs	Xe + $6s$	3,88
56	Barium	Ba	Xe + $6s^2$	5,19
57	Lanthan	La	Xe + $6s^2 5d$	5,59
58	Cerium	Ce	Xe + $6s^2 4f^2$	6,54
59	Praseodymium	Pr	Xe + $6s^2 4f^3$	5,76
60	Neodymium	Nd	Xe + $6s^2 4f^4$	6,31
61	Promethium	Pm	Xe + $6s^2 4f^5$	
62	Samarium	Sm	Xe + $6s^2 4f^6$	6,55
63	Europium	Eu	Xe + $6s^2 4f^7$	5,64
64	Gadolinium	Gd	Xe + $6s^2 4f^7 5d$	6,65
65	Therbium	Tb	Xe + $6s^2 4f^9$	6,74
66	Dysprosium	Dy	Xe + $6s^2 4f^{10}$	6,82
67	Holmium	Ho	Xe + $6s^2 4f^{11}$	
68	Erbium	Er	Xe + $6s^2 4f^{12}$	
69	Thulium	Tm	Xe + $6s^2 4f^{13}$	
70	Ytterbium	Yb	Xe + $6s^2 4f^{14}$	7,06
71	Lutetium	Lu	Xe + $6s^2 4f^{14} 5d$	
72	Hafnium	Hf	Xe + $6s^2 4f^{14} 5d^2$	5,5
73	Tantal	Ta	Xe + $6s^2 4f^{14} 5d^3$	6
74	Wolfram	W	Xe + $6s^2 4f^{14} 5d^4$	7,18
75	Rhenium	Re	Xe + $6s^2 4f^{14} 5d^5$	7,87
76	Osmium	Os	Xe + $6s^2 4f^{14} 5d^6$	8,7
77	Iridium	Ir	Xe + $6s^2 4f^{14} 5d^7$	9,2
78	Platin	Pt	Xe + $6s^2 4f^{14} 5d^8$	8,96
79	Gold	Au	Xe + $6s 4f^{14} 5d^{10}$	9,20
80	Quecksilber	Hg	Xe + $6s^2 4f^{14} 5d^{10}$	10,39
81	Thallium	Tl	Xe + $6s^2 4f^{14} 5d^{10} 6p$	6,08
82	Blei	Pb	Xe + $6s^2 4f^{14} 5d^{10} 6p^2$	7,39
83	Bismut	Bi	Xe + $6s^2 4f^{14} 5d^{10} 6p^3$	8,0
84	Polonium	Po	Xe + $6s^2 4f^{14} 5d^{10} 6p^4$	
85	Astat	At	Xe + $6s^2 4f^{14} 5d^{10} 6p^5$	
86	Radon	Rn	Xe + $6s^2 4f^{14} 5d^{10} 6p^6$	10,69
87	Francium	Fr	Rn + $7s$	
88	Radium	Ra	Rn + $7s^2$	5,21
89	Actinium	Ac	Rn + $7s^2 6d$	
90	Thorium	Th	Rn + $7s^2 6d^2$	
91	Protactinium	Pa	Rn + $7s 5f^2 6d$	
92	Uranium	U	Rn + $7s^2 5f^3 6d$	
93	Neptunium	Np	Rn + $7s^2 5f^4 6d$	
94	Plutonium	Pu	Rn + $7s^2 5f^5 6d$	
95	Americium	Am	Rn + $7s^2 5f^6 6d$	
96	Curium	Cm	Rn + $7s^2 5f^7 6d$	
97	Berkelium	Bk	Rn + $7s^2 5f^8 6d$	
98	Californium	Cf	Rn + $7s^2 5f^{10}$	
99	Einsteinium	Es	Rn + $7s^2 5f^{11}$	
100	Fermium	Fm	Rn + $7s^2 5f^{12}$	
101	Mendelevium	Md	Rn + $7s^2 5f^{13}$	
102	Nobelium	No	Rn + $7s^2 5f^{14}$	
103	Lawrencium	Lw	Rn + $7s^2 5f^{14} 3d$	
104	Kurtschatowium	Ku		
105	Bohrium	Bo		
106				
107				

Konstante	Symbol	Zahlenwert mit Fehlergrenze und Einheit
Gravitationskonstante	γ	$= (6{,}6670 \pm 0{,}007) \cdot 10^{-11}\ \mathrm{m^3\,s^{-2}\,kg^{-1}}$
Normalfallbeschleunigung	g	$= 9{,}8065\ \mathrm{m\,s^{-2}}$
Masseeinheit der physikalischen Masseskala	$1\ ME_{ph}$	= 1/16 der Masse des Isotops O-16
Masseeinheit der chemischen Masseskala	$1\ ME_{ch}$	= 1/32 der Masse des natürlichen O_2
Verhältnis der Masseeinheiten	ME_{ph}/ME_{ch}	$= (1{,}000272 \pm 0{,}000007)$
Masseeinheit der auf C-12 bezogenen Masseskala	$1\ u$	= 1/12 der Masse des Isotops C-12
Verhältnis	u/ME_{ph}	$= 1{,}000317917$
Absolute Masseeinheit	m_u	$= 1{,}660565 \cdot 10^{-27}\ \mathrm{kg}$
Molvolumen idealer Gase unter Normalbedingungen	V_0	$= (22{,}4136 \pm 0{,}0030)\ \mathrm{m^3\,kmol^{-1}}$
Avogadrokonstante	N_A^{ph}	$= (6{,}02252 \pm 0{,}00016) \cdot 10^{26}\ \mathrm{kmol^{-1}}$
	N_A^{ch}	$= (6{,}02336 \pm 0{,}00020) \cdot 10^{26}\ \mathrm{kmol^{-1}}$
Universelle Gaskonstante	R_0^{ph}	$= (8316{,}96 \pm 0{,}34)\ \mathrm{J\,kmol^{-1}\,K^{-1}}$
	R_0^{ch}	$= (8314{,}67 \pm 0{,}40)\ \mathrm{J\,kmol^{-1}\,K^{-1}}$
Boltzmannkonstante	$k_B = R_0/N_A$	$= (1{,}38054 \pm 0{,}00007) \cdot 10^{-23}\ \mathrm{J\,K^{-1}}$
Loschmidtkonstante	$N_L = N_A/V_0$	$= (2{,}68699 \pm 0{,}00010) \cdot 10^{25}\ \mathrm{m^{-3}}$
Permeabilität des Vakuums (Magnetische Feldkonstante, Induktionskonstante)	$\mu_0 = 4\pi \cdot 10^{-7} \times \mathrm{VsA^{-1}m^{-1}}$	$= 1{,}256637061 \cdot 10^{-6}\ \mathrm{Vs\,A^{-1}\,m^{-1}}$
Dielektrizitätskonstante des Vakuums (Elektrische Feldkonstante, Influenzkonstante)	$\varepsilon_0 = 1/\mu_0 c^2$	$= (8{,}854186 \pm 0{,}000012) \cdot 10^{-12}\ \mathrm{As\,V^{-1}\,m^{-1}}$
Elementarladung	e	$= (1{,}60210 \pm 0{,}00003) \cdot 10^{-19}\ \mathrm{As}$
Faradaykonstante	$F^{ph} = N_A^{ph}\, e$	$= (9{,}65219 \pm 0{,}00011) \cdot 10^{7}\ \mathrm{As\,kval^{-1}}$
	$F^{ch} = N_A^{ch}\, e$	$= (9{,}64953 \pm 0{,}00016) \cdot 10^{7}\ \mathrm{Askval^{-1}}$
Vakuumwellenwiderstand	Γ	$= (367{,}730) \pm 0{,}001)\ \Omega$
Lichtgeschwindigkeit im Vakuum	c	$= (2{,}997925 \pm 0{,}000003) \cdot 10^{8}\ \mathrm{m\,s^{-1}}$
Plancksches Wirkungsquantum	h	$= (6{,}6256 \pm 0{,}0005) \cdot 10^{-34}\ \mathrm{Js}$
Drehimpulsquantum	$\hbar = h/2\pi$	$= (1{,}05449 \pm 0{,}00008) \cdot 10^{-34}\ \mathrm{Js}$
Stefan-Boltzmannsche Strahlungskonstante	$\sigma = 2\pi^5 k^4/15c^2h^3$	$= (5{,}6697 \pm 0{,}0010) \cdot 10^{-8}\ \mathrm{W\,m^{-1}\,K^{-4}}$
Wiensche Konstante	K	$= (2{,}8978 \pm 0{,}0004) \cdot 10^{-3}\ \mathrm{mK}$
Plancksche Strahlungskonstanten	$c_1 = 8\pi hc$	$= (4{,}9916 \pm 0{,}0002) \cdot 10^{-24}\ \mathrm{J \cdot m}$
bzw.	$c_1 = hc^2$	$= (5{,}9544 \pm 0{,}0006) \cdot 10^{-17}\ \mathrm{W\,m^2}$
	$c_2 = hc/k$	$= (1{,}43879 \pm 0{,}00002) \cdot 10^{-2}\ \mathrm{m \cdot K}$
Rydbergkonstante	$R_\infty = m_e e^4/8\varepsilon_0 ch^3$	$= (1{,}09737309 \pm 0{,}00000012) \cdot 10^{7}\ \mathrm{m^{-1}}$
Sommerfeldsche Feinstrukturkonstante	$\alpha = e^2/2\varepsilon_0 ch$	$= (7{,}29720 \pm 0{,}00003) \cdot 10^{-3}$
Comptonwellenlänge des Protons	$\lambda_c^p = h/m_p c$	$= (1{,}32140 \pm 0{,}00002) \cdot 10^{-15}\ \mathrm{m}$
Comptonwellenlänge des Neutrons	$\lambda_c^n = h/m_n c$	$= (1{,}31958 \pm 0{,}00005) \cdot 10^{-15}\ \mathrm{m}$
Comptonwellenlänge des Elektrons	$\lambda_c^e = h/m_e\, c$	$= (2{,}42621 \pm 0{,}00005) \cdot 10^{-12}\ \mathrm{m}$
Ruhmasse des Elektrons	m_e	$= (9{,}1091 \pm 0{,}003) \cdot 10^{-31}\ \mathrm{kg}$
Ruhmasse des Protons	m_p	$= (1{,}672523 \pm 0{,}000003) \cdot 10^{-27}\ \mathrm{kg}$
Ruhmasse des Neutrons	m_n	$= (1{,}67482 \pm 0{,}00002) \cdot 10^{-27}\ \mathrm{kg}$
Massewert des Elektrons	m_M^e	$= (5{,}48597 \pm 0{,}00006) \cdot 10^{-4}\ u$
Massewert des Protons	m_M^p	$= (1{,}072863 \pm 0{,}0000002) \cdot u$
Massewert des Neutrons	m_M^n	$= (1{,}0086630 \pm 0{,}0000017)\ u$
Massewert des Wasserstoffatoms	m_M^H	$= (1{,}0078251 \pm 0{,}0000002)\ u$
Massewert des C-12-Atoms	m_M^C	$= (12{,}000000 \pm 0)\ u$
Massewert des O-16-Atoms	m_M^O	$= (15{,}9949150 \pm 0{,}0000006)\ u$
Massenverhältnis Neutron/Proton	m_n/m_p	$= 1{,}001373$
Massenverhältnis	m_p/m_e	$= 1836{,}10 \pm 0{,}02$

Konstante	Symbol	Zahlenwert mit Fehlergrenze und Einheit
Proton/Elektron		
Massenverhältnis Neutrales Pion (Meson)/Elektron	m_{π^0}/m_e	$= 134{,}972$
Massenverhältnis Lambdahyperon/Elektron	m_Λ/m_e	$= 1115{,}544$
Spezifische Ladung des Elektrons	e/m_e	$= (1{,}758796 \pm 0{,}000002) \cdot 10^{11}$ As kg^{-1}
Spezifische Ladung des Protons	e/m_p	$= (9{,}57900 \pm 0{,}00004) \cdot 10^{7}$ As kg^{-1}
Bohrscher Radius (des Wasserstoffatoms)	$r = 4\pi\varepsilon_0\hbar^2/m_e e^2$	$= (5{,}29167 \pm 0{,}00002) \cdot 10^{-11}$ m
Klassischer Elektronenradius	$r_e = e^2/4\pi\varepsilon_0 m_e c^2$	$= (2{,}81777 \pm 0{,}00004) \cdot 10^{-15}$ m
Bohrsches Magneton	$\mu_B = eh/2m_e$	$= (9{,}2732 \pm 0{,}0004) \cdot 10^{-24}$ A m^2
Kernmagneton	$\mu_K = eh/2m_p$	$= (5{,}05050 \pm 0{,}00020) \cdot 10^{-27}$ A m^2
Astronomische Entfernungen		
Astronomische Einheit	1 AE	$= 1{,}4959787 \cdot 10^{11}$ m
Parsek	1 pc	$= 3{,}0856 \cdot 10^{6}$ m
Lichtjahr	1 Lj	$= 9{,}460895 \cdot 10^{15}$ m
Astronomische Zeitmaße:		
Tropisches Jahr	1 a_t	= 365,24219 mittlere Sonnentage
Siderisches Jahr	1 a_s	= 365,25636 mittlere Sonnentage
Sterntag	1 d_{st}	= 23 h 56 min 4,098 s
Geschwindigkeit der Radiowellen	c	$= 2{,}99792458 \cdot 10^{8}$ ms^{-1}
Hubblekonstante	H	$= (75 \pm 25) \cdot 10^{9}$ m · pc · s^{-1}
Unsere Galaxis:		
Sternzahl		10^{11}
Masse		$2{,}7845 \cdot 10^{41}$ kg
Radius		25000 pc
Rotationsgeschwindigkeit (am Sonnenort)		200 km s^{-1}
Mittlerer Sternabstand		1 pc
Sonne:		
Sonnenmasse	m_s	$= 1{,}989 \cdot 10^{30}$ kg
Sonnenradius	r_s	$= 6{,}96 \cdot 10^{8}$ m
Sonnenleuchtkraft (Strahlungsleistung)	P_s	$= 3{,}82 \cdot 10^{26}$ W
Solarkonstante		= 1360 W m^{-2}
Mittlere Dichte	$\bar{\varrho}_s$	$= 1{,}41 \cdot 10^{3}$ kg m^{-3}
Effektive Temperatur	T_{eff}	= 5770 K
Absolute Helligkeit	M	= 4,79 Größenklassen
Scheinbare Helligkeit	m	= −26,8 Größenklassen
Spektraltyp	G2	
Mittlere Entfernung Sonne – Erde	1 AE	$= 1{,}4959787 \cdot 10^{11}$ m

Register

Die Stichwörter sind alphabetisch nach der Buchstabenfolge in ihnen bis zum Komma geordnet; nach dem Komma beginnt ein neues Alphabet. Die Umlaute „ä", „ö", „ü" werden wie „a", „o", „u" eingeordnet.
Griechische Buchstaben und Zeichenfolgen, die nicht nur Buchstaben enthalten, sind nach der Aussprache alphabetisiert worden. Der Buchstabe „ß" ist wie „ss" eingeordnet. Besteht ein Stichwort aus Adjektiv und Substantiv, so wird das Adjektiv in der Regel nachgestellt, z. B. ist das Stichwort „reziprokes Gitter" unter „Gitter, reziprokes" zu finden.
Haben mehrere aufeinanderfolgende Stichwörter gleiche Buchstabenfolge vor dem Komma, so werden sie vom zweiten Stichwort ab durch einen Gedankenstrich ersetzt. Der Gedankenstrich steht auch für den durch einen senkrechten Strich abgetrennten ersten Teil eines vorhergehenden Stichworts.
Ein Zusatz „f." oder „ff." hinter der Seitenzahl weist darauf hin, daß auch die nachfolgende Seite bzw. mehrere nachfolgende Seiten nachzuschlagen sind.

A

Ablenkung im | elektrischen Feld 61
– magnetischen Feld 62
Abschirmung von Neutronen 172
Absorption 188, 587, 589
Absorptions|index 589
–kante 223
–koeffizient 336, 587
–spektralanalyse 220
–spektrum 266
Abweichung, mittlere quadratische 83
Acetylaceton 251
Acetylen 241
Achse, optische 593
–, polare 356
Achsendispersion 594
(8-*N*)-Regel 367
Ad-Atom 467
Additionsreaktion, chemische 438, 444
Addition von Drehimpulsen 100
Additiv 414
Adiabate 321
Adhäsion 404
Adhäsions|arbeit 404
–spannung 412
Ad-Ion 467
$A^{III}B^{V}$-Verbindung 358, 362
Adsorbat 405
Adsorbens 405
Adsorption 379, 405, 454
– (Temperaturabhängigkeit) 408
–, negative 407
Adsorption, physikalische 405
–, selektive 408
–, spezifische 464
Adsorptions|enthalpie, isoster molare 409
–isotherme 408
–komplex 405
–mittel 405
–schicht, monomolekulare 408
–verdrängung 408
Adsorptiv 405
Aerosol 415
Affinität 308
Aggregatzustand 302, 604
–, flüssiger 327
Aktionskonstante 429
Aktivatoren 602
Aktivierung 164
Aktivierungs|analyse 165
–energie 174, 366, 425, 429, 433, 468
–enthalpie 432
–entropie 432
–volumen 434
Aktivität 31, 309
–, optische 252, 595
Aktivitätskoeffizient 309, 344, 420, 447
akustische Resonanz 227
akustischer Zweig der Gitterschwingungen 474
Akzeptor 508
–charakter, Lösungsmittel 435
Albit 374
Alfvénsche Mach-Zahl 663
Alfvénwellen 614
Alkalimetall 367, 547
–halogenid 362
–nitrat 456
Alkalispektrum 224
Allotropie 375
α-Polonium 367
α-Praseodymium 367
Alphazerfall 150
α-Zinn 368
Al-Razi 10
Altersbestimmungen durch Radionuklide 168
Aluminium 367, 388
Alvarez 700
Amaldi 153
Ambarzumjan 722
Ammoniak 449, 451, 453
–synthese 422, 459
Ampere 573
Ampholyte 446
Amplitude 53, 72
Analyse, coulometrische 460
–, thermische 377
Anaximenes 9
Anbonden 533
Andalusit 374
Anderson 40, 472
Andrade 31
Andromedanebel 715
Anharmonizität der Gitterschwingungen 485f.
Anhydrit 456
Anionenelektrode 459
Anisotropie 352, 588, 593, 612
Anlagerung von Elektronen 608
Anlassen 378
Anlaufvorgang 455
Anode 466

Anodenfall 623
-gebiet 624
Anorthit 374
Anorthosit 671
Anregung 608
- von Kernen 140
Anregungs|energie 608
-zustand, kollektiver 505
Anreicherungs|betrieb 528
-schicht 528
Antiferro|elektrika 571
-magnetika 585
-magnetismus 641
Anti|materie 120
-mon 367
-neutron 35
-poden, optische 252
-proton 35
-quark 125
anti-Stokessche Banden 271
Anti|symmetriebedingung 239
-teilchen 35, 116, 120
Apex 713
Aphel 676
Apoenzym 437
Apogäum 668
Apollo-Asteroiden 683
Apsidenlinie 669
Äquipartitionsprinzip 322
Äquivalent|dosis 163
-leitfähigkeit 347
Äquivalenzgesetz, fotochemisches 437
Aquokomplex 453
Arbeit 55
Arbeitsspeicher 586
Arcatomverfahren 627
ARISTOTELES 9
ARRHENIUS 449
Arrheniusgleichung 429
Arsen 367
Artefakten 472
ARTSIMOWITSCH 648
a-Silicium 599
Assoziationskolloide 416
Asteroiden 676, 683
-gürtel 683
Asthenosphäre 655
Astrophysik 646
Asymmetriezentrum 252
Äther|hypothese 9
-theorie des Lichts 38
Atmosphäre, untere 656
- der Erde 656
Atmosphärenschichtung der Sterne 694f.
Atom 9
Atom|abstandsbestimmung 283
- -Atom-Stoß 217
Atombindung 236
Atombindung, polare 243
Atome, wasserstoffähnliche 201
Atom|gitter 361
-hypothese 9
Atomisierungsenthalpie 362, 365
Atomistik, wissenschaftliche 11
Atom|modell, statistisches 208
-radius 368
-rumpf 224
-spektralanalyse 218f.
-spektrum 608
Ätzgruben 387
Aufbauprinzip, Atome 240
Aufdampfen im Hochvakuum 533
Auflösungs|geschwindigkeit 337
-vermögen 24, 78, 181
Aufstäuben im Hochvakuum 533
Aufwachsverfahren 458
Auger|effekt 600
-spektrometer 600
Ausdehnung, thermische 391
Ausdehnungskoeffizient 393
-, kubischer 486
-, linearer 486
-, linearer thermischer 391
-, thermischer 304, 316, 332
Aushärtung 392
Auslösezählrohr 160
Auspuffgas 619
Ausschließungsprinzip 29
Austausch|energie 109, 211
-kräfte zwischen Nukleonen 147
-reaktion 438
-stromdichte 466
Austenit 378
Austrittsarbeit für Metalle 542
Auswahlregeln 28, 220, 261
- für Elektronenübergänge 264
- für Strahlungsübergänge 221
- für Wellenzahlvektoren 476
Automatisierungstechnik 530
Autokatalyse 436, 456
Autoprotolyse 447
-konstante 449, 451
Autoradiogramm 167
Autoradiographie 167
Auxochrome 268
AVOGADRO 12
Avogadrokonstante 14, 318
Avogadrosche Hypothese 12
Avogadrosches Gesetz 318
AXFORD 664

B

BACON 10
Bahn|drehimpuls 52
-moment, magnetisches 574
Balkenspiralen SB 716
BALMER 26, 221
Balmerserie 26, 222, 688
Band|abstand 499, 505, 514, 592, 641
-generator 33
-lücke 475
-paßfilter 593
-struktur 493
-verbiegung 494
-zweig 495
Bändermodell 361, 371, 504f.
- der Kristallelektronen 488
- im Ortsraum 493
- im periodischen Zonenschema 495
- im reduzierten Zonenschema 494
- in drei Dimensionen 495
Bänderüberlappung 493
BARDEEN 524, 525, 555
Barium|carbonat 453
-titanat 569
-titanatkeramik 569f.
BARKHAUSEN 583
Barkhauseneffekt 583
BARKLA 28, 185
BARTLETT 254
Baryonzahl 118, 126
Basalt 671
Base, harte 441
-, weiche 441
Basisschaltung 526
BASOV 191
Bauelement, aktives 530
-, passives 530
Baufehler, makroskopische 383
-, mikroskopische 383
-, submikroskopische 383
Bauteil, optisches 592
Bayard-Alpert-Röhren 632
BCS-Theorie 555
BEAMAN 382
beam-foil-Spektroskopie 218
BEATT 320
BECHER 554
BECKER 34
Bedeckung, multimolekulare 408
Bedeckungsveränderliche 686
BEILBY 387
Beilbyschicht 387, 401
BENEDICT 320
Benetzung 412ff.
Benetzungsgleichgewicht 412

Benetzungsspannung 412
Benitoit 374
Benzen 251, 382
–carbonsäure 445
–dicarbonsäure 448
–spektrum 265
Becquerel 31
BECQUEREL 29f.
BERGMANN 642
BERNOULLI 14
Bernstein 379
Beryll 374
BERZELIUS 12
Beschichten 472
Beschichtungsverfahren 629
Beschleunigung 53
Beschleunigungszahl 728
Besetzung der Elektronenzustände 496
Besetzungszahl 107
–inversion 192
–formalismus 107
Bestandteile, disperse 415
Betatron 33, 115
Betazerfall 149
BET-Gleichung 408
BETHE 647
Bethe-Weizsäcker-Formel 138
Beugungsreflexe 480
Beugung von | Elektronenstrahlen 491
– Röntgenstrahlen 491
Beweglichkeit 503, 612
– der Ladungsträger 517
Bewegung, thermische 609
– im Zentralfeld 59
Bewegungsgleichung 54
– der Kristallelektronen 500f.
– nach DRUDE und LORENTZ 564f.
Bezugs|elektrode 461
–system 45
–zelle 462
–zellspannung 462
Bilanzgleichungen 514
Bildungs|enthalpie 306
–wärmen 306
Bildwandler, thermischer 332
–röhren 184
Bindung, chemische 235
–, kovalente 235, 654
–, metallische 235f., 535, 654
Bindungsenergie 243f.
–, Atomkerne 137
–, Elektronenpaare (im Supraleiter) 553
–, mittlere 365
–, mittlere molare 442
Bindungs|kräfte im Festkörper 485
–kraftkonstante 279
Bindungs|länge 244
–ordnung 240
–verhalten 654
–winkel 241
–winkelbestimmung 283
Binnendruck 319
Binomial|koeffizient 85
–verteilung 85
Biopolymere 298
Bipolar|transistor 521
–technik 519, 530
BIRKELAND-EYDE 628
Bismut 367
BJERRUM 449
Blasenkammer 117
Blei 365, 368
–methode 163
Blitzerscheinung 619
BLOCH 227, 492f.
Bloch|funktion 493
–wand 583
Blochsche Näherung 492
Blochsches Theorem 471, 493
Block|copolymer 294
–zahl 294
BLOEMBERGEN 596
BODENSTEIN 427
Bogenentladung 619, 624
BOHR 27f., 43, 153
Bohrsche Postulate 27
Bohrscher Radius 43, 197
Bohrsches | Magneton 43, 574
– Modell 510
BOISBAUDRAN, DE 13, 25
Bolometer 184, 274
BOLTZMANN 14, 307
Boltzmann|faktor 429, 432
–konstante 89, 244, 484
–statistik 483, 498, 540, 605
–verteilung 92
Bonden 527
Bor 367
Borat 374
BORELIUS 537
Bor|hydrid 254
–kammer 160
BORN 28, 95, 237, 362, 472, 485
Bornitrid 643
–, kubisches 643
Born-Oppenheimer-Näherung 431
Bortrifluorid 247
BOSCH 459
Bose | -Einstein-Statistik 483, 497, 552
–kondensation 109
–verteilung 92
Bosonen 100, 107, 555
–, intermediäre 133
BOTHE 34
Bourdonmanometer 635
BOYLE 10
Boyle-Mariottesches Gesetz 10, 318
Boylesches Gesetz 632
Boyletemperatur 319
BRACKETT 26, 32
Brackettserie 222
BRAGG, W. H. 358
BRAGG, W. L. 358
Braggsche | Gleichung 359, 480
– Kurve 158
BRAHE 694
BRATTAIN 525
BRAUN 421
Braunstein 459
Bravaisgitter 352f.
Brechung 587
– von Teilchenstrahlen 75
Brechungs|gesetz 74
–index 73, 558, 614
–index, komplexer 589
Brechzahl 587f.
BRECKENRIDGE 569
Breitlinienspektroskopie 230
Brekzien 670
Brems|strahl 159
–strahlung 28, 608, 723
–verhältnis 176
–vermögen 176
Brennspannung 622
Brewsterwinkel 185, 596
BRIDGMAN 320, 357, 642
BRIEGLEB 369
Brillant 643
BRILLOUIN 491
Brillouinsche Näherung 491
Brillouin|streuung 191
–zone 371, 474, 479, 491f., 544, 548
BROGLIE, DE 28
BRÖNSTED 446
BROWN 13
Brownsche Bewegung 13
Bruchgrenze 388
BRUNAUER 408
Brutreaktor 650
–, schneller 179
Brüter 175
Brutstoffe 176
Brutto|reaktion 419
–reaktionsgleichung 419
–stabilitätskonstante 452
Bubblespeicher 587
BUDKER 112
BUERGER 638
Bullvalen 251
BUNSEN 25, 218f.
BURGERS 385
Burgersvektor 389f.

Bursters 707
Büschelentladung 619
Butanol 250

C

Cadmiumsulfid 506
Caesium 605
–chlorid 372, 455
CAHN 252
Calcitstruktur 372
Calcium|carbonat 455
–oxalat 456
Canning 178
CANNIZZARO 12
Carbanionen 444
Carbeniumionen 444
Carbonado 642
Carbonsäure 435
Catenane 256
CCD-Technik 529
C-Gitter 353
Cementit 371, 378
Centaurus A 720
Čerenkov|effekt 117
–zähler 117
Cermets 539
CHADWICK 34, 156
Chalkogenidglas 379, 381, 506
CHANDRASEKHAR 705
Chandrasekhargrenze 705
charge-transfer-Banden 245
Charm 119, 123, 126
Chemilumineszenz 437
Chemisorption 405, 409, 458
Chemisorptionsenthalpie 409
Chelat 285
–polymer 285
Chemolumineszenz 266, 600
CHESLEY 569
Childsches Gesetz 566
–, modifiziertes 566f.
Chiralität 252
Chlorknallgaszelle 461
CHOCHLOW 596
Chromophor 268
Chromosphäre 672
cis-trans-Isomerie 252, 297
CLAUSIUS 14, 307, 563
Clausius-Clapeyronsche Gleichung 333
CLUSIUS 326
Cluster 330, 385
C-MOS-Technik 529
Cobalt 367
COCKROFT 33, 111
Coenzym 437
Colour 125
Compoundkernmodell 153
COMPTON 40
Compton|effekt 39, 159
–wellenlänge 40
Computer 522, 530
COOPER 555
Cooperpaar 505, 555ff.
Copernicanische Idee 723
Copeumlagerung 251
Copolymer, linearer 294
–, statistisches 294
Corioliskraft 656
Coulomb|anregung 137
–barriere 151
–integral 238
–potential 362f.
–wall 150
Coulombsches Gesetz 237, 614
Coulombwechselwirkung 65, 614
– in der Atomhülle 210
Coulometer 460
COUPER 235
COWAN 36
CRICK 300
Cr_3Si-Struktur 550
Cristobalit 372, 456f.
CROMER 369
CROOKES 15, 24
CT-Banden 245
Cuban 256
Curie 31
CURIE, M. 29f.
–, P. 29f., 580
Curie|konstante 568
–temperatur 641
– -Weiß-Gesetz 581, 568
Cyclo|addition 440
–alkan 253
–butan 256
–polymerisation 293
–propan 256
Czochralskiverfahren 357

D

DALTON 11f.
Daltonsches Partialdruckgesetz 318
Dampf 317
–, übersättigter 375
Dampfdruck 332
–diagramm 338
–erniedrigung 338
– idealer Mischungen 338
–kurve 374f.
– über gekrümmten Oberflächen 402
Daniellelement 461
Daten, kritische 320
Dauermagnet 581
DAVISSON 41
DDL-Logik 531
DEBIERNE 30
DE BOER 458
DE BROGLIE 40, 93, 95
de Brogliesche Beziehung 41, 490
DEBYE 245, 290, 343, 364, 480, 483, 484, 487
Debyegrenzfrequenz 552
Debye-Hückel-|Onsager-Gleichung 348
– Theorie 343
Debye|funktion 485
–länge 515
–näherung 484
Debyescher Radius 344
Debye-Scherrer-Verfahren 360
Debyesches T^3-Gesetz 485
Debyetemperatur 481
–, charakteristische 485
Debye-Waller-Faktor 481
Defekte, eindimensionale 385
–, nulldimensionale 385
–, zweidimensionale 385
Defektelektron 502, 547
Deformation, elastische 388
–, plastische 388
De-Haas-van-Alphen-Effekt 546
Dehnbarkeit, reversible 394
d-Elektron 236
δ-Orbital 239
Dembereffekt 600
DEMOKRIT 9ff.
DERJAGIN 330
Desaktivierung 189
DESCARTES 37
Desolvatation 417
Desorption 405, 454
Destillation 339
–, isotherme 289
Destriaueffekt 601
Detektor 524
Deuterium 33, 646
Deuteron 146
Diacetyldioxim 452
Diagnostik, optische 617
–, spektroskopische 617
Dialyse 415
Diamant 367, 642
–herstellung 628
–körper, polykristalliner 642
–modifikation, hexagonale 643
–produktion 642
–struktur 640
–synthese 358, 365, 367, 422, 635, 641
–typ 371

Diastereoisomere 252
Dichroismus 185, 595f.
Dichte, kritische 725
–verteilung, radiale 315
–welle 714
Dichtungen bei Druckkammern 634
DICKEL 326
Dielektrika|eigenschaften 560f.
–stoffe 560f.
Dielektrikum 560
Dielektrizitätskonstante 451, 589
–, absolute 570
–, differentielle 570
–, komplex relative 561
–, relative 561
Diels-Alder-Reaktionen 440
Diethylether 250
Differenzierungsprozeß, magmatischer 671
Differenzthermoanalyse 381
Diffusion 325f., 345, 370, 383, 390, 434, 468
–, ambipolare 516, 612
–, gebundene 392
Diffusions|koeffizient 326, 346, 391, 515
–längen 599
–grenzstromdichte 468
–mechanismus 390
–schicht 337, 468
–spannung 461, 520
–stromdichte 515
–strom in Halbleitern 504
–überspannung 467f.
Dilatanz 418
Dimethyl|formamid 435
–sulfoxid 435
Diopsid 374
Diode 521
Dioden-Transistor-Logik 531
Dipol, induzierter 561
–, permanenter 561
Dipolmoment 370, 562
– (chemische Bindung) 245
–, elektrisches 64
–, induziertes 576
–, magnetisches 64
–bestimmung 284
DIRAC 35, 203, 498, 637
Dispersion 415, 589f.
–, anormale 590
–, normale 590
Dispersions|beziehung 72
–eigenschaft 614
–energie 245
–mittel 415
–relation 473
–theorie 590
Dispersoidkolloid 416
Dissoziation 443, 608
– des Eisens 703
Dissoziations|energie 258
–enthalpie 243
–grad 420f.
–konstante (Base) 447
–konstante (Säure) 443, 445, 447
DOEBEREINER 12, 458
Domäne 569
–, ferromagnetische 583
Domänen|struktur 570
–wand 569, 583
Donator 508
–charakter eines Lösungsmittels 435
Donnanspannung 461
Doppel|bindungsregel 367
–brechung 185
–helix 300
–resonanz, akustische 233
Doppelschicht, elektrochemische 464
–kapazität 465
Doppel|spaltexperiment 37
–spirale 300
–stern 686
–strangpolymer 295
Doppler|effekt 75, 482, 617, 689
–radar 618
–verbreiterung 190, 617
–verschiebung 618
DORN 30
Dosimetrie 162
Dosis, letale 167
Dosisleistung 163
Dotant 509
Dotierung 358, 383, 455, 506, 508f.
DOT-Speicher 587
Draht, stromdurchflossener 65
DRAPER 437
Drehachse 247, 354
Drehimpuls|eigenwerte 98
–operator 97
–satz 52
Drehkristallmethode 360
Drehmoment 57
–, elektrischer Dipol 65
–, magnetischer Dipol 67
Drehschieberpumpe 629f.
Drehspiegelachse 248, 354
Drehspiegelung 248, 354
Drehung 247, 354
Dreierstoßrekombination 607
3-Kelvin-Strahlung 709, 724, 729
Dreiphasengrenzlinie 410, 412
Dreizentrenbindung 254
DRICKHAMER 641, 642
Drift|bewegung 612
–geschwindigkeit 503, 612
–kammer 117
–strom 515
Drosseleffekt, isenthalpischer 321
Druck 311
–, dynamischer 633
–, hydrostatischer 632f.
–, osmotischer 340
–, quasihydrostatischer 633
–, statischer 633
Druckerzeugung 633
–, quasihydrostatische 634f.
Druck|fixpunkt 636
–homologieregel 639
–ionisation 620
–koeffizient 464
–koordinationsregel 639
–messung 635ff.
– Tiefen-Charakterisierung 645
–vervielfacher 634
–waage 636
–wasserreaktor 178f.
DRUDE 564
DSCHABIR 10
DUHEM 305
Dulong-Petitsche Regel 483, 540
DUNGEY 664
Dunkel|entladung 624
–wolke 708, 711
DUOUING 596
Duplexfilme 411
Durchbruchspannung 527
Durchgang von | α-Teilchen 158
– β-Strahlen 159
– γ-Strahlen 158
Durchlaßrichtung 520
Durchschlag, elektrischer 518, 567
–, elektromechanischer 567
–, thermischer 567
Durchschlags|kanal 567
–vorgang 567
Durchtritts|faktor 468
–reaktion 467
–überspannung 467f.
Duroplaste 392
Dynamik der|Hochatmosphäre 659
– Kristallelektronen 499
Dynamo|mechanismus 661
–theorie 654f.
Dystektikum 377

E

Ebullioskopie 289
Echelettegitter 182
Edelgas 628
-konfiguration 205, 235, 237, 255
-verbindung 236, 254
Edelsteine, synthetische 644
Edukte 419
Effekt, anharmonischer 485, 590, 596
-, elektrophoretischer 348
-, magnetoakustischer 546
-, thermoelektrischer 504, 537
-, thermomagnetischer 504
Effektivpotential 59, 104, 198
Effusion 323
EHRENBERG 165
EHRENFEST 381
Ehrenfestsche Sätze 110
Eigen|drehimpuls 52
-fehlordnung 385, 454
-frequenz 58, 79
-funktion 97
-leitung 506f., 511
Eigenschaften, extensive 303
-, intensive 303
Eigenschwingung 193
Eigenwert 97
-, entarteter 492
-aufspaltung 492
-gleichung 97
Eindringtiefe, Londonsche 551
Einelektronennäherung 488
Einfangfaktor 525
Einfriertemperatur 380, 392f.
Einkomponentensysteme 374
Einkristall 357, 359, 395, 471
E1-Reaktion 444
E(k)-Verlauf 495
Einlagerungsmischkristalle 370
Einschränkungen 374
Einschubmethode 621
Eisen 388
- -Kohlenstoff-Diagramm 378
-nickel 654
EINSTEIN 9, 13, 38f., 187, 191, 482, 484, 726, 727
Einstein|beziehung 516
-funktion 484
-koeffizient 188f., 220f.
-näherung 484
Einstein-de-Sitter-Kosmos 723f.
Einsteinsche Formel 51
Eis 246, 372, 375
-essig 451
Eiweißstoffe 294
Elaste 392
Elastizitäts|gesetz 477
-modul 394
Elastomere 392
ELDOR 233
Elektrode 459
-, einfache 460
-, mehrfache 466
- zweiter Art 459, 463
Elektroden|polarisation 466
-potential 462
-spannung, relative 462
Elektro|fotografie 507, 598, 603
-kristallisation 469
Elektrolyse 461
-, Grundgesetze 15
Elektrolyte 341, 640
-, echte 341
-, mittelstarke 341
-, potentielle 341
-, schwache 341
-, starke 341
Elektrolytlösung 341
Elektron 15, 120
Elektronegativität 244
Elektron-Elektron-Wechselwirkung 497, 503
Elektron|-Loch-Flüssigkeit 505
- -Loch-Paar 505, 518
- -Kern-Doppelresonanz 232
- -Neutrino 120
- -Phonon-Streuung 476, 517
- -Phonon-Wechselwirkung 497, 503, 552, 555, 592
- -Photon-Vertex 128
- -Photon-Wechselwirkung 514
- -Positron-Paar 489
- -Positron-Vernichtung 131
Elektronen, äquivalente 209
-, freie 663
-, nichtäquivalente 209
Elektronenaffinität 237, 244, 362, 608
- der spezifischen Wärmekapazität 554
Elektronen|anlagerung 608
Elektronenanregung (in Ionenkristallen) 601
- von Molekülen 260
Elektronenanregungs|bande 263
-spektrum 264
Elektronenanteil der | spezifischen Wärmekapazität 485, 541
- Wärmeleitung 487, 504
Elektronen|austrittsarbeit 498
-besetzung 498
-beugung 40, 359, 360f
Elektronen|delokalisierung 368
-dichte 237
- -Donator-Akzeptor-Komplex 245
-doppelresonanz 233
-einfang 149
-emission 542
- -Energie-Verlust-Analyse 505
Elektronengas 236
-, entartetes 491, 497
Elektronen | im Potentialkasten 490
-konfiguration 29, 204
-konzentration 541
-lawine 605, 621
-leitung 488
-mangelverbindung 254
-mikroskop 20
-niveaustruktur des Leiters 547
Elektronenpaar 553
-akzeptor 245
-donator 245
Elektronen|polarisation 562, 591
-produktion, maximale 660
-quelle 542
-radius, klassischer 16
Elektronenresonanz, ferromagnetische 230
-, paramagnetische 443
Elektronen|röhre 530
-schall 614
-schwingung 613
-spin 16, 574
-spinresonanz 228, 576
Elektronenstoß 606
-ionisation 17
-ionisierung 605
-ionisierungskoeffizient 605
Elektronen|strahleinschuß 605
-strahlen 480, 650
Elektronenstrahl|lithographie 532f.
-röhre 20
-sonde 618
Elektronen|streuung, elastische 130
-temperatur 610
-transport, ballistischer 529
Elektronenübergänge, optische 496
-, optisch induzierte 481
Elektronen|verteilung, radial 201
-volt 65
Elektronenzustand, lokalisierter 494

Elektronenzustand, stationärer 489
– im Festkörper 488
– im periodischen Potential 491
Elektrostriktion 573
Elektrostriktionskoeffizient 573
Element, galvanisches 461
Elementar|ladung 15f., 63
–prozeß im Plasma 606
–reaktion 426
–teilchen 115ff.
–teilchenphysik 35ff.
Elementhalbleiter 384
Eliminierungen (chemische Reaktion) 438
ELSASSER 41
Emaniermethode 165
Emission, spontane 188
–, stimulierte 188, 710
–, thermische 17f.
Emissions|nebel 708
–spektralanalyse 219
–spektrum 266, 270
–theorie des Lichts 37
Emitter|wirkungsgrad 525
–schaltung 525
EMK 464
EMMETT 408
EMPEDOKLES 9
Emulgator 417
Emulsion 415, 417
Emulsionspolymerisation 417
Enantiomere 252
enantiotrop 375
Endgruppen 293
–bestimmung 290
ENDOR-Technik 232
Energie, freie 307, 311
–, innere 305, 310, 332
–, kinetische 48, 50
–, potentielle 56
–, relativistische 49
–, thermische 89
Energie|abstrahlung, infrarote (Strahlenchemie) 659
–ausbeute 438
–bänder 492
Energiedichte, elektrische 67
–, magnetische 67
Energieeinheit 71
–, atomare 43, 197
Energie|einstrahlung 658
–erhaltung 476
–erhaltungssatz 14, 49, 305
–formen 70
Energielücke 492
– eines Supraleiters 552
Energie|operator 97
–quelle, alternative 646
Energiesatz 70
–, mechanischer 56
Energie|schale 205
–spektrum des Wasserstoffatoms 198
–strom 70
–stromdichte 70
–verteilungssatz, Boltzmannscher 429
–zustände von Molekülen 256
ENSKOG 325
Entartung 98
–, zufällige 105, 199
– des Elektronengases 498
Entartungsgrad beim Wasserstoffatom 199
Entelektrisierungs|faktor 563
–feld 563
Entfernung, fotometrische 728
Enthalpie 305, 311
–, freie 307, 311, 422f., 639
Entladung, selbständige 622
–, unselbständige 622
Entladungsmechanismus 621
Entropie 82, 307, 310
–satz 306
–elastizität 394
– idealer Gase 314
Enzym 299, 437
Epitaxie 358
–schicht 522
–technik 522, 527, 533
EPR 228
Erdalkalimetallchalkogenide 362
Erdbeben 654
–wellen 652
Erde (Planet) 652
Erd|inneres 653
–kern 652, 654
–körper 652
–kruste 652
–magnetfeld 655
–mantel 652
erdnaher Raum 652
E-Region 660
Erhaltung | des Isospins 126
– der Parität 126
Erhaltungsgrößen, Quantentheorie 109
Erhaltungssatz, Baryonen und Leptonen 724
–, Drehimpuls 126
–, Energie und Impuls 126
Erinnerungseffekt 395
Erosionserscheinung 679
Erscheinung, seismische 655
–, tektonische 655
–, vulkanologische 655
Eruptionsveränderliche 693
ESAKI 522
Essigsäure 449
Esterspaltung 436
Ether 435
Ethylen 241f.
–diamintetraessigsäure 452
E_T-Werte 435
Eutektikum 376
EVANS 432
Evolution des Kosmos 724f.
Ewaldkugel 79, 480
Ewaldsche Konstruktion 480
EWING 580
Exosphärentemperatur 656
Expansion der Welt 717
Explosivschweißen 645
Extinktion, interstellare 711f.
Extinktionskoeffizient 269
Extraktion 512
Extrusion 644
Exziton 505, 592, 598
EYRING 329, 429
Eyringgleichung 430f., 434
E-Z-Nomenklatur 253

F

Fabry-Perot|-Interferometer 182
– -Resonator 193
FAIRFIELD 665
Faktor, präexponentieller 433
–, präexponentieller (Reaktionskinetik) 429
–, stöchiometrischer 419
Fällungs|chromatographie 287
–gleichgewicht 452
–indikator 453
–reaktion 453
–titration 292, 453
Faltblattstrukturen 299
FARADAY 15
Faraday|effekt 618, 706, 711f.
–konstante 460
Faradaysche Gesetze 460
Farbbildröhre 602
Farbe 267
Farben|-Helligkeits-Diagramm 691, 704
–index 688
Farb|erscheinung 269
–stoff 267f.
–temperatur 687, 711
Färbung von Kristallen 591
Farbzentrum 383
Fehlordnung, chemische 358, 383, 509

Fehlordnung, elektrische 383
–, nulldimensionale 509
–, strukturelle 383ff.
–, thermodynamische 385
Feinstruktur 26f., 202, 211, 225
– der Spektrallinien 225
–konstante 203
Feld, dynamo-elektrisches 661
–, elektrisches 64
–, elektromagnetisches 68
–, erdmagnetisches 661
–, magnetisches 65
–, quasistationäres 62
–, statisches 62
–begriff 62
Feldeffekt 527
–transistor 525, 527
Feldemission 17f., 543, 566
Feldemissionsstromdichte 543
Feld|energie 67
–ionisation 18
–spat 374
Feldstärke, kritische 550
–, magnetische 66
f-Elektronen 236
Ferment 437
FERMI 34f., 153, 208, 498, 637
Fermi-Dirac-|Statistik 483, 497, 511, 544
–Verteilungsfunktion 507, 540
Fermienergie 92, 496, 498, 520, 541
Fermifläche 544, 591
–, offene 545
Fermigas, entartetes 92
Fermikörper 545, 547
–hälse 546
–querschnitte, extremale 546
Ferminiveau 505, 507, 511, 522
Fermion 100, 107, 496
Fermische Grenzenergie 92
Fermiverteilung 92
Fernordnung 366, 379
Fernsehaufnahmeröhre 598
Fernwirkungstheorie 68
Ferrimagnetismus 585, 641
Ferrit 378, 585
–kernspeicher 586
Ferrocen 255
Ferro|elektrika 567f.
–magnetika 580f.
–magnetismus 582, 641
–spinell 585
Festkörper, idealer 302, 471
–, perfekter 471
–, realer 498
Festkörper|aktivierung 383
–chemie 471
Festkörper|diagnostik 543
–effekt 233
–eigenschaften, optische 587f.
–optik 587
–physik 471
–reaktivität 370, 383
–schaltkreis 530
Feststoffreaktion 454f.
Feynmangraph 127
F-Gitter 353, 367
Fibrillarkolloid 416
Fibrillen 396
Ficksches | Diffusionsgesetz 326
– Gesetz 515
Film, gasanaloger 411
–, kondensierter 411
–bildung 398
–druck 411
FISCHER, E. 437
Fission 646
Fixstern 619
FIZEAU 38
Flächen|defekte 385, 387
–diode 521
–satz 57
–speicher 586
–transistor 525
Flammenerscheinung 619
Flash 703
Fleckenrelativzahl 674
FLEROW 151
Fließ|grenze 388
–zone 396
–zonenverfahren 357f.
Flocculi 673
Flockung 416
Flotation 408
Fluchtgeschwindigkeit 717
Flugzeit, mittlere freie 517, 539
Fluidität 345
Fluoreszenz 266, 270
–spektrum 270
Fluß, magnetischer 67
Flüssigkeit 316, 327
–, autophobe 411
–, kristalline 327, 330
Flüssigkeitsoberflächenstruktur 400
Flüssigkristallanzeige 331
Flußschlauch 550
FOCK 204
Folgereaktion 428
Folie, orientierte 397
Fontäneeffekt 349
Form, kritische 151
Formalladung 236
Formfaktoren, elektromagnetische 130
Foto|chemie 437
–diode 525
–dissoziation 656
Fotoeffekt 159, 597
–, äußerer 38
Foto|element 598
– -EMK 525, 598
Fotografie 602
–, Grundprozesse 588
Fotoionisierung 17, 605, 606
Fotoleitung 597
–, primäre 598
–, sekundäre 598
Foto|maske 532
–maskierung 526
–resist 532
–spannung 598
–transistor 526, 599
–zelle 184
FOUCAULT 38
Fourier|integral 73
–raum 479
–transformation 73
$f\tau$-Wert 150
Fragmentierungsphase 700
Fraktionierung 287
FRANCK 26f., 217
Franck-Condon-Prinzip 264
FRANKEN 596
Frank-Hertz-Versuch 26
FRANKLAND 235
FRANKLIN 14, 411
Fransenmodell 395
FRAUNHOFER 25, 38, 182, 219, 672
F-Region 660
Freiheiten 374
Freiheits|grad 374
–zahl 501
FRENKEL 385, 524
Frequenz 53, 72
–faktor 429
–messung (Radiowellenbereich) 183
–standard 195
–verdopplung 596
–verdreifachung 597
–verteilung 58
–zählung 183
FRESNEL 38, 589
Fresnelkoeffizienten 590
FREUNDLICH 408
Freundlichgleichung 408
Friedmannalter 727
FRIEDRICH 358
FRISCH 34
FRÖHLICH 555
Fugazität 308
Fumarsäure 448, 450
Funken|entladung 624
–erosion 628

Fusion 646
Fusions|hybridreaktor 650
–reaktor 556
F-Zentrum 383, 601
F′-Zentrum 384

G

galaktisches Zentrum 714
Galaxis 712, 715
–, aktive 714
–, irreguläre 716
–, quasistellare 722
–entstehung 725
–kern 712f.
–masse 718
–rotation, differentielle 713
Gallium|arsenid 357, 362, 495, 506
–phosphid 357, 507
Galvanische Zelle 461
Galvanispannung 460
Gammaquant 147
Gantmachereffekt 546
Gas 317
–, ideales 311f.
–, interstellares 708f.
–, permanentes 317
–, reales 315, 317, 319
Gasentladung, elektrische 621
Gasentladungs|lichtquelle 625
–plasma 620
–röhre 621
–schalter 627
–schaltröhre 627
Gas|detektor 530
–gemisch 318
–konstante, molare 318, 484
–nebel, leuchtende 710
–reaktion, homogene 314
GASSENDI 11, 235
Gas|theorie, kinetische 14, 322
–zustand, idealer 302, 317
Gateisolation 560
Gaußverteilung 85
GAY-LUSSAC 12
Gay-Lussacsches Gesetz 12, 318, 320
Gefrierpunktserniedrigung 339
Gegenwirkungsprinzip 54
GEIGER 24, 31, 216
Geiger-Müller-Zählrohr 160
Geißlerröhre 626
Gel 417
GELLMANN 36
Gelpermeationschromatographie 288
Gemisch, aceotropes 339, 377
Generationsrate 599
Generation von Ladungsträgern 512
Generator, magnetohydrodynamischer 628
Genfer Nomenklatur 250
Gerdienbogen 624
Germanium 357, 362, 368, 384, 386, 506
GERMER 41
Gesamtdrehimpuls 52
–, Elektronenhülle 209
Gesamtenergie 56
–, Molekül 261
Gesamtwirkungsquerschnitt 152
Geschwindigkeit 47, 53
–, mittlere thermische 609
Geschwindigkeits|konstante 425, 431f.
–verteilungsfunktion 617
–verteilungsgesetz 610
Gesetz der | konstanten Proportionen 10, 11
– multiplen Proportionen 11
– Partialdrücke 12
– unabhängigen Ionenwanderung 347
Gesetz des idealen Gaszustands 318
Gesetz von | CLAUSIUS und MOSSOTTI 563
– HESS 306
– KIRCHHOFF 306
Gestaltelastizität 394
Getterpumpe 631
Gezeitenschwingung 659
g-Faktor 213
GIBBS 305, 307, 374
Gibbs- | Duhemsche Gleichung 304
–Helmholtz-Gleichung 423
Gibbssche|Adsorptionsgleichung 405
– Fundamentalgleichung für die Grenzflächenphase 399
Gibbssches Phasengesetz 374
GILLESPIE 248
GINZBURG 568
Gips 456
Gitter, eindimensionales 352
–, reziprokes 478, 480
–, zweidimensionales 352
–beweglichkeit 517
–defekt 503
–dynamik 471
gitterdynamische Theorie 485
Gitterenergie 361ff.
Gitterschwingung 472
–, akustische 477
Gitterschwingung, optische 474
Gitter|sonde 617
–spektrograph 78
Gittervektor, reziproker 479, 487
Glanzwinkel 360
Glas 379
–, metallisches 379, 382
–, technisches 644
–bildung 379ff.
–bildungstendenz 381
GLASSTONE 429
Glaszustand 392
Gleichgewicht, chemisches 309, 419f.
–, dynamisches 425
–, energetisches 697
–, gekoppeltes 421
–, mechanisches 697
–, radioaktives 32, 157
–, thermisches 14
Gleichgewichtsbedingung, thermodynamische 308
Gleichgewichtsgalvanispannung 460
Gleichgewichtskonstante 422f., 425, 429, 431f.
–, konventionelle 420
–, thermodynamische 309, 420
Gleichgewichts|plasma 615
–reaktion 427
–zellspannung 464
–zustand, thermodynamischer 303
Gleichung, chemische 419ff.
– von GIBBS und HELMHOLTZ 307
Gleichverteilungssatz 91, 539
– der statischen Mechanik 483
Gleit|spiegelebene 356
–spiegelung 356
Gleitung 388
Glimmentladung 379, 624
Glimmentladungs|gleichrichter 627
–lichtquelle 626
Glimmlicht, negatives 623
Glühemission bei Zimmertemperatur 543
Glühkatode 542
Glühkatoden|gleichrichter 627
–ionisationsmanometer 632
Glühlampe 625
Gluonen 36, 133
Glyceraldehyd 252
Golayzelle 184
GOLDBERG 15

GOLDHABER 156
GOLDMAN 569
GOLDSTEIN 15
GOUDSMID 16
GOUY 13
GPC 288
GRAAFF, VAN DE 33
GRAHAM 415
Graham-Bunsensches Gesetz 323
Granat, ferrimagnetischer 585
Granit-Gneis-Schicht 654
Granulation 672
Granulen 696, 672
Graphit 367, 378, 500
–, interkaliertes 500
Graphoepitaxie 533
Gravitations|kollaps 700
–potential 726
–rotverschiebung 686
–theorie 706
GREINACHER 111
Grenzflächen 398
–, gekrümmte 401
–aktivität 405
–energie 399
–enthalpie, spezifische freie 399
–frequenz 484
–spannung 403
GRIMALDI 37
Größe, molare 303
–, spezifische 303
Großer Roter Fleck 680f.
Große Unifikation 134
Großresonanzen 144
Großwinkelkorngrenze 387
GROTTHUS 437
Grundschwingung 80
GRÜNEISEN 488
Grüneisen|beziehung 488
–konstante 488
–parameter 638
Gruppen|geschwindigkeit 73, 500
–schwingung, charakteristische 276
Guinierverfahren 360
Gunneffekt 519
GURNEY 603
Gußeisen 378
GUTENBERG 653
Gutenbergsche Zone 655
Güteschaltung eines Lasers 194
Gyre 354
Gyroide 356
gyromagnetischer Faktor 64, 213
gyromagnetisches Verhältnis 226

H

HABER 459
Hadronen 122
Haftspannung 412
Haftstellen 566, 602f.
– niveau 566
HAHN 34, 154, 165
Halbleiter 488, 492, 497, 499, 505, 640
–, amorphe 511
–, direkte 496, 600
–, entartete 512
–, indirekte 496, 598, 600
–, optische Eigenschaften 592
–, organische 506f.
–, schmalbandige 507, 598
–bauelement 506, 519
–detektor 162, 602
–empfänger 184
–gleichrichter 523
–laser 512, 601
–technik 530
Halbwertszeit 31, 156f., 426
HALL 516, 642
Hall|effekt 516, 600
–koeffizient 517
–leitfähigkeit 661
HALLWACHS 38
Halogenidgläser 379
Halopopulation 715
Halterung, magnetische 647f.
– eines Plasmas 613
Hamiltonoperator 97, 237, 242
HANNAY 642
Härte 366
Hartelastomere 395
Härten 378
HARTREE 204
Haupt|dielektrizitätskonstante 594
–feld, erdmagnetisches 655
–gruppenmetalle 367
–quantenzahl 199
–reihenstern 690f., 699
Hauptsatz, erster 305
–, zweiter 306
Haupt|schnitt 595
–spannung 633
–trägheitsachse 256
Hedvallregel 456
HEIKKILA 665
HEISENBERG 28, 34, 43, 147
Heisenbergsche Unschärferelation 363, 490
Heißpressen, isostatisches 644
Heliosphäre 685
Helium|methode 168
–reaktion 698
Helix 299
Helixstruktur 298, 301
HELLER 554
Helligkeit, absolute 717f.
–, scheinbare 717
– eines Sterns 687
HELMHOLTZ 14, 15, 307
Helmholtzfläche, äußere 464
–, innere 464
Helmholtzschicht, äußere 464
Hemmung, kinetische 437
Henry-Daltonsches Gesetz 336
HERAKLIT 9
HERTZ, G. 26f., 217, 323, 333
HERTZ, H. 15, 38, 69
Hertzsprung|lücke 691
– -Russell-Diagramm 690, 701, 704f.
HERZFELD 429
HESS 306
Hetero|epitaxi 529
–lyse 442f.
–sphäre 656
–übergänge 512
HEVESY 164
Hill-Baldes-Methode 292
Himmelsblau 269
HINES 664
HINSHELWOOD 430
HIPPEL 569
Hirayama-Familien 683
HMO-Methode 243
Hochatmosphäre 656f.
Hochdruck|experimente 654
–forschung, geophysikalische 645
–lampe 626
–modifikation 549, 638
–oxid 654
–physik 632ff.
–plasma 615
–synthese 641
Hochfrequenz|entladung 625
–sonde 617
–spektroskopie 227
–transistor 527
–verhalten 527
Hochgeschwindigkeitskinematographie 618
Hochspannungsleuchtröhre 626
Höchstdruck|lampe 626
–reaktion, dynamische 486
Höchststromentladung 624
HOFMEISTER 701
Hohl|katodenlampe 219
–leiter 281
Hohlraum|resonator 80, 281
–strahlung 42, 186
Hologramm 195

Holographie 195
Homo|epitaxi 529
–lyse 442
–polymer 293
–sphäre 656
Hooke 10, 388, 477
Hoppingleitung 504, 565
Horizont 728
HSAB-Konzept (Säure-Base-Reaktion) 441
Hubble 715
Hubble|alter 727
–beziehung 717, 721
–konstante 717
Hückel 242, 343
Hückelregel 242
Hughes 435
Hugoniot|kurve 638
–zustandsgleichung 638
Hume-Rothery-Phasen 371, 548
Hundsche Regeln 211, 240
Huygens 37f.
Hybridisierung 240
Hybridisierungstendenz 367
Hybrid|schaltung 527
–technik 527
Hydratation 341f., 370
Hydratations|enthalpie 343
–zahl 343
Hydrathülle, äußere 342
–, innere 342
Hydratisomerie 251
Hydro|formulierung 437
–lyse 438, 449
Hydronen 36
Hydroniumionen 343, 446
Hydro|philierung 408
–phobierung 408
–sol 415
–thermalsynthese 358, 643
Hyperfeinstruktur 26f., 214, 225, 229, 282
–messung 284
Hyperladung 119, 122, 126
Hyperonen 36
Hysteresekurve 570, 580

I

Identität der Teilchen 107
Identitätsoperation 247
I-Effekt 445
I-Gitter 353, 367
Ignitron 627
Impaktkrater 670
Impfkristall 357, 375
Implantieren 472
Impuls 48
Impuls, relativistischer 49
– des Kristallelektrons 490
–echolotung 660
–erhaltung 476
–erhaltungssatz 48
–laser 194
–operator 96
–plasma 615
–sättigungsmethode 227
Impulsverteilung 91
–, quantenmechanische 95
Indexellipsoid 594
Indikatormethode 164
Indiumantimonid 357, 362
Indizes, Millersche 353
INDOR 232
Induktions|gesetz 68
–periode 456
Industrierobotertechnik 530
Inertialsystem 45
Infrarot|detektor 598
–optik 591
Infrarotspektroskopie 272
–, Anwendung 276
–, interferometrische 273
Infrarot|spektrum 276
–strahlung 714
–strahlungsquelle 272, 700f.
Ingold 252, 435
Inhibitoren 436
Injektion 598
– neutraler Wasserstoffstrahlen 649
Injektionslumineszenz 600
Injektion von Ladungsträgern 512
Inkrement 424
Inkrementsystem, thermodynamisches 424
Inselstruktur 412
Instabilität 614f.
Integrationsgrad 531
Interferenz 37, 290
–erscheinung 76
–komparator 183
– von Wahrscheinlichkeiten 94
Interferometer 182
–, optisches 617
Interkombinationsverbot 224f.
INTOR 650
intracavity-Spektroskopie 220
Inversion 248, 354
Inversions|drehachse 354
–drehung 354
–konzentration 508
–schicht 524, 527
–temperatur 321
–zentrum 354
Iod 375
Ionen 15, 341
Ionen, mehrfach geladene 607
–äquivalentleitfähigkeit 347
–beweglichkeit 347
–beziehung 235f.
–charakter (chemische Bindung) 244
–implantation 24, 509, 521, 529, 629
–gitter 361f., 383
–kristall 638
–kristall (optische Eigenschaften) 590
–leitung 488, 565
–modell 361
–polarisation 561
–produkt des Wassers 447
–quelle 18
–radius 244, 368
–reibung 659
–schall 613f.
–schwingung 613
–spektrum 608
–strahlen 533, 650
–stoßionisation 17
–temperatur 610
–wanderung, unabhängige 347
–wertigkeit 237
–wolke 344
–zerstäuberpumpe 631
–zerstäubung 649
Ionisation 443, 604, 606
–, spezifische 158
–, thermische 18, 605, 620
–, thermosphärische 660
Ionisations|grad 615, 620, 694
–kammer 117, 160
–stufe 607
Ionisierung der Störstellen 511
Ionisierungs|energie 237, 242, 362, 604
–potential 244
Ionomer 285, 392
Ionosphäre 619, 659, 661
–, Mars 680
–, Venus 678
Ionosphären|plasma 659
–schicht 614
–störung 676
IR-Mikroskop 592
Irrflug 85
irreversibel 304
IR-Spektroskopie 240
IR-Spektrum 641
isobar 304
Isobare 134
Isobarenregel 136, 139
isochor 304
Isochronzyklotron 113
Isolator 488, 492, 497, 499

Isolierstoff 488
Isomer 140
Isomerie 250
–, optische 252
Isomerisierung (chemische Reaktion) 438
Isomerisierungsreaktion 438
Isospin 119
isotherm 304
Isotone 134
Isotope 16, 134
–, radioaktive 32
Isotopen|häufigkeit 690
–trennung 136
–verdünnungsmethode 165
Isotopieeffekt 136, 225, 552
Isotropie 611
–verschiebung 215
IWANENKO 34

J

JANSSEN 25
JEANS 187
Jetstream 683
JOFFÉ 524
JOLIOT, F. 32, 34
JOLIOT-CURIE, I. 32, 34
JOSEPHSON 556
Josephson|effekt 533, 556
–element 559f.
–kontakt 559
–logik 560
–tunnelelement 558
JOULE 14
Joulesche Wärme 65
Joule-Thomson|-Effekt 321
– -Koeffizient, differentieller 321
Jupiter 680f.
–monde 681f.

K

Käfigeffekt 435
Kalium 653
–chlorat 458
–hydrogensulfat 597
Kalkspat 595
Kalomelelektrode 463
Kalorimeter 117
Kalottenmodell 369
Kälte|maschine 351
–mischung 377
Kaltkatodenionisationsmanometer 632
Kaltpressen, isostatisches 644
KAMERLINGH-ONNES 548
Kanalstrahlen 15
Kanalstrahlentladung 624
Kaonen 36
–, neutrale 123, 130
Kapazitätsdiode 522
Kapillar|depression 403
–druck 402
Kapillare 402
Kapillar|kondensation 402
–konstante 403
KAPIZA 349
KARMAN, VON 472, 485
Kaskadengenerator 111
Katalysator 436, 458
Katalyse 379
–, heterogene 458
–, homogene 436
Katode 466
Katodenfall 622
–raum 623
Katodenstrahlung 15
Katodolumineszenz 266, 600
Kautschukelastizität 394
KAUZMANN 382
Kavitation 334
KEESOM 244, 364
Keim|bildung 455
–bildungsgeschwindigkeit 380
–wachstum 603
KEKULÉ 235
Kelvingleichung 402
KEPLER 694
Keplersches Gesetz 676
Kernbrennstoff 176
Kerne, deformierte 137
Kernenergie 646
Kernfusion 35, 629, 646
–, Sonne 672
Kerninduktion 231
Kern-Kern-Doppelresonanz 232
Kern|kräfte 145
–kollaps 146
–kraftwerk 179
Kernladungszahl 134
–, effektive 206
Kern|magneton 43, 141, 575
–modell 141
–moment, magnetisches 575
–plasma 615
–polarisation, dynamische 233
–quadrupolresonanz 234
–radius 136
Kernreaktion 152ff.
– mit Alphateilchen 155
– mit Deuteronen 155
– mit Neutronen 154
– mit Protonen 155
– mit schweren Ionen 155
Kernreaktion mit Teilchen höchster Energie 156
Kernreaktor 34
Kernresonanz, hochauflösende 230
–, paramagnetische 230
Kernrotation 144
Kernspaltung 34, 71
–, spontane 150
Kernspin 141
–resonanz 230
Kern|spuremulsion 162
–umwandlung 32
KERST 33
Keto-Enol-Tautometrie 420
Ketten|isomerie 250
–längenverteilung 286
–reaktion 34, 174, 428
Kieselglas 380
Kinetik, elektrochemische 459, 465
– chemischer Reaktionen 425
Kingscher Ofen 620
KIPPENHAHN 701
KIRCHHOFF 25, 42, 218, 272, 306
Kirchhoff-Planck-Funktion 695
Kirchhoffsches Gesetz 186
Kirkwoodlücke 683
KLEBER 383
Kleine Planeten 683f.
Kleinwinkelkorngrenze 387
Knight-Shift 230
KNIPPING 358
Knotenlinie 669
KNUDSEN 333
Knudsengebiet 326
Knudsensche Effusionsmethode 323
Koagulation 416
Koagulationsstruktur 417
Koaleszenz 416
Koerzitivfeldstärke 570, 581
Koexistenzbedingung 309
Kohärenz 77
–, Laserstrahlung 77, 193
–länge 77, 596
Kohäsion 404
Kohäsions|arbeit 404
–druck 319
–energie 391
–kräfte 391
Kohlelichtbogen 625
Kohlendioxidmolekül 659
Kohlenstoff|detonation 702
– -Stickstoff-Zyklus 698f.
–zyklus 647
Kohlenwasserstoff 435
KOHLRAUSCH 347, 348
Kohnanomalie 485

KOLBE 235
Kollektivbeschleuniger 112
kollektive Erscheinungen 615
Kollektorschaltung 526
Kollisionszone 655
Kolloide 415
–, anisometrische 416
–, hydrophile 416
–, hydrophobe 416
–, isometrische 416
–, lyophile 416
–, lyophobe 416
Koma 684
Kombinationsprinzip 26
Kometen 684
–kern 684f.
–schweif 684
–wolke 684
Komplementarität 41
Komplex, aktivierter 429f., 431
Komplex|bildungsfunktion 449
–gleichgewichte 452
Komplexometrie 452
Komponenten 374
Kompressibilität 363, 366, 376
Kompressibilitäts|faktor 320
–koeffizient 304, 316, 332
Kompressions|manometer 632
–modul 388
–verhalten, Festkörper 638
Kompressor 633
Kondensation 334
Kondensations|geschwindigkeit 333f.
–struktur 417
Konduktionspumpe, elektromagnetische 628
Konfiguration 250, 369
–, Makromoleküle 296
Konfigurationsisomerie 251
Konformation 250, 369
–, Makromoleküle 297
–, Proteine 299
Konformations|elastizität 394
–isomerie 253
Konformere 253
KÖNIG 14
Konstitution 250
–, chemische 314
–, Makromoleküle 292
Konstitutions|formel 241
–isomerie 250
Kontakt|potentialdifferenz 613
–spannung 461
–vergiftung 458
–winkel 412
Kontinentaldrift 654
Konvektion 654, 699
Konversion, innere 148
Konversions|elektronen 148
–koeffizient 148
Konzentration 304, 420
Konzentrationsüberspannung 469
Koordinationsisomerie 251
Kopplung, optoelektronische 599
–, schwache (Supraleiter) 557
Korngrenzen 387, 391, 511
–diffusion 454
Korona 673
–entladung 619
–spektrum 673
Korrelation 83
Korrespondenzprinzip 28, 320
Korund 357
Kosmologie 725f.
–, relativistische 726f.
kosmologisches | Glied 727
– Prinzip 723
Kosmos 724f.
–forschung 358
KOSSEL 224, 235, 469
Kovalenz 236
Kovolumen 319
Kraft 54
–, elektromotorische 464
–, konservative 56
– -Dehnungs-Diagramm 396
–gesetz 54f.
Krater, positiver 625
k-Raum 479, 489, 544
KRAUSE 656
Kreis|bewegung 53
–beschleuniger 112
–frequenz 53, 72
–nebelpulsar 702, 706
–prozeß 361, 363, 365
–strom 575
Kristall, eindimensionaler 477
–, optisch einachsiger 595
–, piezoelektrischer 572
–, pyroelektrischer 573
–, Realstruktur 357, 382f.
Kristallchemie 383
–, energetische 361
Kristall|defekt 471
–detektor 282
–elektron 488f.
–feldaufspaltung 229
–fotoeffekt 600
–gemisch 53
–gitter 471
–gitterspektrometer 182
Kristallisation 380
–, fraktionierte 377
Kristallisationsenthalpie 357
Kristallisations|inhibitor 469
–überspannung 467, 469
Kristall|klassen 352, 354, 356
–physik 383
–phosphor 601
–struktur, Metalle 534
–strukturanalyse 358f., 370
–system 352, 391
–wachstumsgeschwindigkeit 380
–züchtung 357
kritische Größe 176
– Menge 176
kritisches Volumen 176
Kronenether 256
Kronig-Penney-Modell 543
Krümmungsdruck 402
Kryo|pumpe 632, 648
–speicher 560
–technik 487
Kugel|packung, dichteste 366f., 372, 388
–sternhaufen 712, 715
Kühlung, thermoelektrische 537
Kupfer 388, 390, 539
– (I)-Oxidgleichrichter 524
KURLBAUM 186
KURTSCHATOW 647
Kurz|kanaltechnik 528
–schluß, elektrischer 619
–wellenfunkverkehr 614

L

Ladung, effektive 370
–, elektrische 62
–, spezifische 15
Ladungs|dichte 63
–einheit 63
–erhaltungssatz 63
Ladungsträger|beweglichkeit 515, 518
–dichte 507
–optik, Hauptsatz 19
–transport 366
–trennung 612, 661
Ladungs|übergang, partieller 467
–unabhängigkeit der Kernkräfte 146
LAIDLER 429
LAMB 203
Lambert-Beersches Gesetz 269
Lambertsches Gesetz 186
Lambshift 203
Lamellenmodell 395
Laminarkolloid 416
LANDAU 349, 706

LANDÉ 369
Landésche Intervallregel 212
Landéscher g-Faktor 213
LANDSBERG 274
Längen|meßtechnik 195
–standard 195
LANGEVIN 575
LANGMUIR 408, 613
Langmuir|isotherme 408
–sonde 616
Lanthan 367
LARMOR 575
Larmor|frequenz 226
–präzession 57, 67, 576
Laser 191, 593, 596
–, Anwendung 195f.
–diode 601
– -Fluoreszenz-Spektroskopie 618
–licht 77
–mikrospektralanalyse 196
–parameter 195
–resonator 193
–spektroskopie 195
–strahlen 650
– -Ultrakurzzeit-Spektroskopie 196
LAUE, VON 358, 554
Lauediagramm 359f.
Lavesphase 371
LAVOISIER 10
Lawinen|durchbruch 518, 523
– -Laufzeit-Diode 518
LAWRENCE 33
LAWSON 647
Lawsonkriterium 647
LCAO-MO-Methode 239, 242
LEBEDEV 70, 369
LE BEL 235
Lebensdauer, mittlere 86, 513
– von Elementarteilchen 118
LE CHATELIER 421
Ledeburit 378
LEE 150
LEED 361
Leer|laufspannung 520
–stellendiffusion 454
Legierung, ferromagnetische 382
–, metallische 53
Leidenfrostsches Phänomen 326
Leistung 55
Leistungs|gleichrichter 521, 523
–transistor 527
Leiter 499
–, metallische 492, 497
–polymere 295
Leitfähigkeit 539
–, elektrische 379, 382, 488, 499, 590, 611, 654
–, elektrische (in Festkörpern) 481
–, elektrolytische 346
–, metallische 544
–, spezifische 346, 488
– von Halbleitern 505
Leitfähigkeits|koeffizient 348
–tensor 661
Leitungsband 495, 499, 505, 547
–, eindimensionales 536
Leitvermögen, elektrolytisches 341
LENARD 15, 24, 38
Lennard-Jones-Potential 322
LEONARDO DA VINCI 10
Lepton 120
–, schweres 122
–zahl 119, 126
Leucht|diode 600
–elektron 29, 205, 224
–kondensator 601
–kraft eines Sterns 687
–kraftklasse 688
–stoff 601f., 626
–ziffer 602
LEWIS 235, 429, 439, 446
Lewis|base 439
–säure 439
Lichtbogen 624
–ofen 627
LICHTENECKER 538
Licht|energie 70
–kurve (Sterne) 692
–leiterkabel 593
–leiternachrichtenübertragung 593
–mischung 183
–quant 39, 70, 187, 606
–quantenhypothese 38
–quantenstoß 606
–streuung 290, 589
–technik 625
–theorie, magnetische 69
–theorie, elektromagnetische 38
–verstärker 597
–verstärkung 191
–weg 728
Ligandenfeldtheorie 268
LINDEMANN 430
Lindemannsche Regel 640
Lindeverfahren 321
Linear|beschleuniger 111ff.
–kolloid 416
Linienbreite 189
–, natürliche 189
Liniendefekt 385
Linien|formfunktion 189
–spektrum 26
–strahlung 608
Linse, elektrostatische 19
–, magnetische 19
Lintheorie 714
Liquidus|kurve 377
–temperatur 381
Lithiumfluorid 370
Lithosphäre 655
Lobatschewskische Geometrie 726
Löcher 209, 384, 502, 506
–leitung 502
–modell 316
–theorie 329
Lokale Gruppe 722
Lokaler Superhaufen 723
LONDON 245, 364
LONDON, F. 554
LONDON, H. 554
London|gleichung 554
–theorie 554
–wechselwirkung 363
LORENTZ 24, 236, 563, 565
Lorentz|kraft 501
–profil 58, 189
–zahl 536
LOSCHMIDT 14
Löslichkeits|koeffizient 336
–produkt 421, 453
Lösungen, ideale 337f.
–, kolloide 415
–, reale 337f.
Lösungs|enthalpie 337
–kältemaschine 351
–temperatur 336
Lösungsmittel, apolare 435
–, dipolar-aprotische 435
–, protische 435
–einteilung 435
–polarität 435
–retention 398
LOWRY 446
Lumineszenz 266, 597, 600
–diode 525
Luminophor 626
LUMMER 186
Lummer-Gehrke-Platte 182
Lymanserie 222
Lyo|gel 418
–sol 415
–sphäre 416

M

Madelung|faktor 363
–konstante 362
Magellansche Wolken 715

magische Zahlen 142
Magnet, supraleitender 556
Magnetfalle 613
–, lineare 648
Magnetfeld 655
–, interplanetares 663
–, interstellares 711f.
–, Sonne 671
–, Stern 688
–sonde 619
magnetische|Kernmomente 141
– Monopole 134
magnetisches Moment von Elementarteilchen 118
Magnetisierung 573
–, Dauermagnet 583
–, spontane 580, 582
Magneto|enzephalogramm 560
–kardiogramm 560
–meter, supraleitendes 560
–plasma 615, 663
–pause 663
Magnetosphäre 663, 665, 677, 681
–, Saturn 682
Magneto|sphärenphysik 663
–striktion 584
–widerstandseffekt 537
Magnetschallwelle 614
Magnetschweif 663, 681
Magnetspeicher 586f.
Magnonen 584
MAJORANA 147
Majoritätsträger 508
–konzentration 511
Makromolekül, Aufbau 284
–, Eigenschaften 284
–größe 286
Maleinsäure 448f.
MALUS 38
MANDELSTAM 274
Manganinmanometer 636
Maniperm 583
MANOW 596
Marangonieffekt 412
MARCELIN 429
MARCKWALD 16, 30
Marefläche 670
MARIOTTE 10
Markariangalaxie 716
Mars 678f.
MARSDEN 24, 216
Mars|kanäle 679
–oberfläche 679
–satelliten 680
Martensit 378
Mascons 670
Masereffekt 710
Masse 48, 117
Masse, effektive 491, 497, 501, 507, 510, 544, 547
–, geschwindigkeitsabhängige 44
–, reduzierte 61
–einheit, atomare 137
Masse-Leuchtkraft-|Beziehung 691f.
– Verhältnis 718
Massen|spektrograph 16, 24
–spektrometrie 619
–trenner 24
–wirkungsgesetz 309, 420, 508, 516, 620
–zahl 134
Massen|bruch 420
–defekt 137
–erhaltungssatz 51
–formeln 124
MASSON 15
Materialeigenschaften im Erdinneren 654
Materialgröße, optische 588
Materie, entartete 698
–, intergalaktische 724
–, interstellare 619, 700, 708ff.
–welle 40, 488
MATTHIAS 549
Matthiessensche Regel 536
MAUGUIN 248
MAXWELL 14, 38, 188
Maxwell|-Boltzmann-Statistik 539
–gleichung 516
Maxwellsche Geschwindigkeitsverteilung 91
Maxwellverteilung 610
MCLEOD 632
MC MILLAN 114
Medium, interplanetares 685
M-Effekt 445
Mehrfach|bindung 241
–ionisierung 607
Mehrquantenprozeß 190
Mehrschichtadsorption 408
Meißner-Ochsenfeld-Effekt 551, 554f., 559
Meißnerphase des supraleitenden Zustands 550
MEITNER 34
MENDELEJEW 12f.
Merkur 677
Mesomerie 241
Meson 36, 122
Meso|pause 656
–sphäre 655f., 659, 662
Meßmethode, kernphysikalische 619
Meßwert, mittlerer 98
–, quantenmechanischer 94, 98
Metagalaxis 724
Metall|bearbeitung 627
–carbonyl 255
–cluster 255
Metalle 488, 638, 640
–, optische Eigenschaften 591
–, amorphe 504
–, edle 463
–, lineare 535
–, unedle 463
Metall|gitter 361, 370
–indikator 453
–ionenelektrode 459
–spitzendiode 524
Metallurgie 627
Meteorit 684f.
Meteoroid 685
Meterdefinition 182
Metgläser 504
Methanmolekül 240
Methode des selbstkonsistenten Felds 106, 204
Metrik 726
MEYER, J. R. 14
MEYER, L. 12f.
MHD-Generator 556, 628
MICHELSON 38, 182, 194
Michelsoninterferometer 182
Mie-Grüneisen-Gleichung 638
Migration 341, 346, 468
Mikro-Brownsche-Bewegung 393
Mikroelektronik 504, 506, 509, 519, 527, 530
–trends 533
Mikro|feld 611
–kryoelektronik 533
–prozessor 522
–strukturierung 533
Mikrotron 112
Mikrowellen|diagnostik 618
–interferometer 618
–radiometer 618
–rauschen 618
–spektroskopie 227, 281, 283
–strahlung 558
Milchstraßensystem 686, 712
–, Entwicklung 714f.
MILLIKAN 16, 39
Millikanversuch 16
Minilöcher 708
Minoritätsträger 508
–konzentration 511
Mischbarkeit 376f.
–, vollständige 335
Misch|katalysator 458
–kristall 383, 538
Mischphase 302
–, flüssige 334
Mischung 302
–, Flüssigkeiten 316
–, ideale 335, 337f.

Mischung, reale 337f.
Mischungs|energie, innere 317
–enthalpie 317, 335
–entropie 317, 335, 424
–lücke 336
–temperatur, kritische 336
MIS-Transistor 527
Mitbewegung des Kerns 202, 215, 222
Mitfällung 453
Mittellinie, optische 594
Mittelwert 83
–, quantenmechanischer 98
Mizellen 406
Mizellkolloid 416
Modell, optisches 154
–, statisches 153
– des idealen Gases 322
– freier Elektronen 491, 540
–potential 493
– unabhängiger Elektronen 203
– unabhängiger Teilchen 106
Moden 186, 188, 193
–synchronisation 194
Moderator 170, 175f.
Modifikation 375
Modifikationswechsel 455
Mohorovičiodiskontinuität 652f.
Moisson 642
Molalität 303, 420
Molekularbiologie 531
Molekular|destillation 333
–strahltechnik 529
Molekularität, chemische Kinetik 440
–, chemische Reaktionen 426
Molekülbanden 608
Moleküle, angeregte 256
–, interstellare 710
Molekülionen, negative 663
Molekül|kolloide 416
–kristalle 591, 638
–modelle 322
Molekülorbital 239, 243
–schema 240
Molekül|plasma 615
–rotation 256
–schwingung 258
–spektrum 256, 608
–struktur 247
–symmetrie 247
Molenbruch 303, 420
Mol|masse 419
–volumen 13, 419
–wärme 305
Moment, magnetisches 574
Mond 668ff.
Mond, physikalische Eigenschaften 669f.
Mond|bahn 668f.
–beben 669f.
–durchmesser 669
–entstehung 671
–gebirge 670
–gestein 670f.
–kern 670
–kruste 669
–mantel 669
–oberfläche 670
–rückseite 670
–systeme 677
–vulkanismus 670
Monochromator 273
monotrop 375
MOS-Bauelement 519
Moseley 13, 28f.
Moseleysches Gesetz 28f.
MOS-Kurzkanal-Transistor 518
Mössbauer 481, 482
Mößbauer|effekt 481
–spektroskopie 482
Mossotti 563
MOS-|Technik 527, 530
– Transistor 525, 527
Mott 504, 524, 566, 603
Mottmodell 566
Mulliken 244
Multiphononenprozeß 481, 514
Multiplett 212
Multiplizität 212, 225
Murnaghan 637
Myon 36, 120
Myonenatome 202
Myonium 202
Myon-Neutrino 120

N

Nachleuchten 601
NaCl-Struktur 362
Näherung freier Elektronen 488, 491
Nahordnung 328, 379
Nahwirkungstheorie 68
Naphthalen 375
Natriumchlorid 362f.
Naturphilosophie, antike 9
–, griechische 9
Nebel, elliptische 715
–, planetarische 710f.
–haufen 722f.
–kammer 30
–leuchten 710
–linien 710
Nebel|veränderliche 701
–zählung 723
Nebengruppenmetalle 367
Nebenquantenzahl 199
Néeltemperatur 641
Neemann 36
Neodymium 367
Nephelin 374
Neptun 682
–monde 683
Nernst 307, 468
Nernst-Noyes-Whitney-Gleichung 337
Nernstsche Gleichung 460
Netz|ebene 78, 359f.
–werkbildner 381
Neuhaus 639
Neukurve 580
Neumann-Koppsche Regel 483
Neutralgas|atmosphäre 656
–wind 661
Neutralschicht 665
Neutralisation 449
Neutralisationsenthalpie 448
Neutralitätsbedingung 511
Neutralteilchentemperatur 610
Neutrino 35, 117, 149, 699f.
–nachweis 121
–oszillation 122
Neutron 33f., 122
Neutronen, ultrakalte 173
–absorber 177
–beugung 170f., 359, 361
–dotierung 509
–erzeugung 168f.
– -Gamma-Karottage 166
–leiter 171
–nachweis 171
–physik 168f.
–quellen 169
–spektroskopie 171
–stern 686, 694, 702, 705
–strahlen 480
–streuung (Anwendung) 171
–überschuß 136
Neutron-Phonon-Streuung 476
Newlands 12
Newmanprojektion 253
Newton 11, 37f.
Newtonsche Bewegungsgleichung 501
Newtonsches | Gravitationsgesetz 676
– Fließgesetz 325
N-Galaxin 716
Nichtbenetzung 413
nichtbindende Zustände 239
Nichterhaltung der Parität 150

Nicht|gleichgewichtsplasma 615
-stöchiometrie 384
-unterscheidbarkeit 483
Nicolsches Prisma 595
Niederdruckplasma 615
NILSON 13
Niveaubreite 140
Nivellierung 448
n-Kanal-MOS-Transistor 528
n-Leiter 384
n-Leitung 506, 508
NMR 230
Norit 671
Normalen|achse, optische 594
-flächen 594
-geschwindigkeit 594
-richtung 594
Normal|leiter 550
-schwingung 260, 473
Normierungs|bedingung 83, 95
-integral 109
Novae 693
N-S-Gesetz 723
Nukleinsäuremolekülstruktur 300
Nukleon 34, 122, 134
Nullfeldaufspaltung 579
Nullpunkts|bewegung 472, 475
-energie 352, 362f.
-energie beim harmonischen Oszillator 103
-entropie 307

O

Oberfläche 398, 472, 511
Oberflächen|diffusion 454
-druck 411
-eigenschaften reiner Stoffe 399f.
-energie 388f.
-energien fester Stoffe 401
-fester Stoffe 400
-film 410ff.
-potentialdifferenz 460
-rekombination 514, 607
Oberflächenspannung 399
-, kritische (der Benetzung) 414
- reiner Flüssigkeiten 400
- von Festkörpern 401
Oberflächenwellen 652
Oberflächenzustand 513
- der Elektronen 524, 528
Oberschwingung 80
Objektmenge 419
OERSTEDT 66, 573
Ohmsche Kontakte 524
Ohmsches Gesetz 461, 503, 515, 518, 536, 612
Olberssches Paradoxon 726
Öldiffusionspumpe 631
Oligomer 286
Öl-in-Wasser-Emulsion 417
Olivin 372
Opazität 695
OPPENHEIMER 237
Optik, integrierte 593
-, nichtlineare 195, 593, 596
optisches Pumpen 233
Optoelektronik 504, 525, 597
Optronik 593, 597
OPW-Methode 493, 543
Orbitalradius 370
Orbittheorie 667
Orientierung, homogene 331
-, homöotrope 331
Orientierungs|polarisation 561
-quantenzahl 100, 103, 199
Orionarm 713
Orthohelium 211
Orthokieselsäure 372
Ortsvektor 53
Osmometrie 289
Osmose 340
OSTWALD 375, 458
Ostwaldsche Stufenregel 642
Oszillator, harmonischer 54, 90, 102, 104, 475, 485
-, isotroper harmonischer 60
OVERHAUSER 233
Overhausereffekt 233
Oxydation 439
Oxydations|mittel 463
-stufe 236
-zahl 236, 439
Ozon 367, 658f.

P

Paarbildung 159
-, innere 148
Paarerzeugung 40, 120
Paarungsenergie 139
Paarvernichtung 40, 120
Packung, hexagonal dicht 367
-, kubisch dicht 367
Palladium 370f.
PANETH 164
PARACELSUS 10
Parahelium 211
Parallelreaktion 428
Paramagnetismus 576ff.
Parameter, kritischer 320
Parität 119
-, Nichteinhaltung 129
Paritätsverletzung in der Atomhülle 133
Partialdruck 319, 420
Partonen 132
PAUL 35
PAULI 106, 149
PAULING 242, 244, 361
Pauliprinzip 29, 106, 240, 248, 489, 496, 540
PASCHEN 186
Paschen-Back-Effekt 214
Paschenserie 222
PEARSON 439, 441, 446
PEIERLS 487
Pedersenleitfähigkeit 661
Pekuliarbewegung 713, 717
Peltier|effekt 537
-wärme 464
Penumbra 673
PENZIAS 724
Peptisation 416
Perigäum 668
Perihel 676
Perioden-Helligkeits-Beziehung 693, 702
Periodensystem der Elemente 12, 206
Peritektikum 377
Perlit 378
Permanentmagnet 581
Permeabilität 589
-, relative 573
-, Vakuum 573
PERRIN 13, 24
Perseusarm 714
PESHKOV 349
Petrologie 645
PETTENKOFER 12
Pfropfpolymer 295
PFUND 26
Pfundserie 222
P-Gitter 353, 367
Phantompolymer 293
Phase 72, 302, 374
-, amorphe 472
-, cholesterische 331
-, metastabile 472
-, nematische 331
-, quasihomogene 302
-, reine 302
-, smektische 331
-, teilamorphe 472
Phasen|anpassung 596f.
-diagramm 374
-geschwindigkeit 72
-gleichgewicht 309, 374
Phasengrenze 398
-, reguläre 414
Phasengrenzfläche 302

Phasengrenzschicht 398
Phasen|kolloid 416
–lage (eindimensionaler Quantenzustand) 557
Phasenumwandlung 483
–, polymorphe 638
*p*H-Indikator 450
Phononen 472, 475f., 483, 486
–, akustische 478
–, optische 478, 587, 590
–anteil der Wärmeleitung 487
–gas 487
Phonon|-Phonon-Streuung 476, 487
– -Photon-Streuung 476
Phosphat 374
Phosphor, schwarz 367, 375
–, weiß 367, 375
Phosphoreszenz 266, 270, 597, 600
Photonen 39, 118, 476, 606
–stoß 606
–streuung, Kristall 479
Photon-Phonon|-Streuung 487
– -Wechselwirkung 481
Photon-Photon-Wechselwirkung 596
Photo|sphäre 672, 694
–synthese 271, 438
*p*H-Wert 447, 449f.
Physisorption 405, 458
PIDDINGTON 664
Piezo|effekt 572f.
–elektrika 573
–elektrizität 568, 571f.
–keramik 572
–modul 572
π-Komplex 254, 437
Pinch|effekt 624, 648
–entladung 624
pin-Übergang 522, 527
Pion 36
Pion-Nukleon-Vertex 128
π-Orbital 239, 242
Piranimanometer 632
Pirouetteneffekt 705
p-Kanal-MOS-Transistor 528
PLANCK 27, 39, 42, 93, 187, 307
Plancksche | Beziehung 101
– Formel 91
– Konstante 28, 42, 475
Plancksches | Strahlungsgesetz 187
– Wirkungsquantum 37
Planeten, jupiterartige 677
–, terrestrische 677
Planetenatmosphäre 677f.
Planeten|beschaffenheit 677ff.
–kosmogenie 685
Planetensystem 676
–, Entstehung 685f.
Planet Erde 652
Planetoid 676, 683
Plasma, isothermisches 615
–, Magnetosphäre 666
Plasma|ätzen 532
–, Erdinneres 620
–brenner 628
–chemie 628
–diagnostik 616
Plasmaelektronen 605
–eigenfrequenz 614
Plasma|erzeugung 619ff.
–instabilität 648
–kanonen 628
–klassifizierung 615f.
–mantel 666
–parameter 615f.
–periode 725
–potential 616
–randschicht 612
–schicht 666f.
–schwingung 675
–sphäre 666f.
–spritzen 628
–strahl 627
–temperatur 615
–vorkommen 619ff.
– -Wand-Wechselwirkung 649
Plasmazustand 604
–, Anwendungen 625ff.
–, Eigenschaften 609ff.
Plasmonen 505
–frequenz 592
Platin 458
PLATON 9
Platzwechsel|mechanismus 565
–vorgang 391
p-Leitung 384, 506, 508
PLÜCKER 15
Pluto 682
pn-Übergang 509, 512, 519f., 525, 598
POHL 601
Poissonsche|Gleichung 321
– Konstante 388
Poissonscher Koeffizient 321
Poissonverteilung 85
POLANYI 429, 432
polare Hörner 666
–bänder 667
Polarimeter 595
Polarisation 364, 466, 561, 568, 612
–, chemisch induzierte 233
–, elektrische 590
–, remanente 570
Polarisation | des Lichts 185
– elektromagnetischer Wellen 69
Polarisationsfilter 595f.
Polarisator 595
Polarisierbarkeit 245, 364
Polaritonen 590
Polarlicht 676
Polarographie 468
Polaroidfolie 596
Polaronen 504, 566
Pole der galvanischen Zelle 461
Poly|addition 439
–elektrolyt 299
–kondensation 439
Polymer, amorphes 298
–, anorganisches 285
–, elementorganisches 285
–, kristallines 298
–, lebendes 288
–, organisches 382
–, orientiertes 396
–, regelmäßiges 293
–, unvernetztes 392
–, vernetztes 295, 392
–, verzweigtes 295
Polymerisation 439, 444
Polymerisationsgrad 285
–, mittlerer 289
–bestimmung 289, 292
–verteilung 286
Polymorphie 375
Polymorphismus 534
Polywasser 330
POMERANTSCHUK 350
Pomerantčukeffekt 350
PONTECORVO 700
Poole-Frenkel-Effekt 566
Population, Sterne 690, 712
positive Säule 623
Positron 35, 40, 503
Positronen|annihilation 546
–emission 149
Positronium 202
Potential 56
–, chemisches 92, 308, 498
–, elektrisches 65
–, elektrochemisches 460
–, inneres elektrisches 462
–fläche 56
–kasten 102
–topf 102, 489
–schwelle 431
Poyntingvektor 70
Präresonanzramaneffekt 275
PRELOG 252
Presse, hydraulische 634
Pressen, isostatische 644
Prezipitation 668
PRINGSHEIM 186

PROCHOROV 191
Proportional|kammer 117
-zählrohr 160
Protein 298f.
Protogalaxis 722
Protolysegleichgewicht 446
Proton 122
Protonen|akzeptor 436, 446
-beugung 361
-donator 436, 446
Protonierungsgrad 450
Proton-Proton-Kette 698f.
Protostern 700
Protuberanzen 674
PROUST 10
PROUT 12
Prozeß, fotografischer 438
Pseudoimpuls 480, 501
– des Kristallelektrons 495
Pseudo|ordnung, Reaktionskinetik 426
-potential 493
-vektor 46
Pufferlösung 450
Pulsar 705f., 712
Pulsationsveränderliche 692f., 702
Pulverdiffraktometer 360
Pumpmechanismen 192
Punkt, kritischer 320, 375
-defekt 385, 391, 509
-gruppe 248, 356
-symmetrieelement 247
-symmetriewelle 652
Pyroelektrizität 573
pyrophor 458

Q

Quadratwurzelgesetz 347
Quadrupelpunkt 376
Quadrupolresonanz 227
Qualitätsfaktor 163
Quanten|ausbeute 437, 597
-bedingung 28
-charakter des Lichts 588
-chromodynamik 132f.
Quanteneffekt 471
– der Gravitation 727
Quanten|elektrodynamik 128
-elektronik 191
-empfänger 274
-flüssigkeit 349
-gas, ideales 91
-hypothese, Plancksche 42
– -Interferenz-Bauelement, supraleitendes 559
-mechanik 235, 471
-physik 9
Quantensprünge 27
Quantentheorie 475
– der Gitterschwingung 475
– der kovalenten chemischen Bindung 492
– der Molwärme 324
Quanten|zahl 98, 489
-zustand, makroskopischer 555
Quarks 36, 125
Quarz 358, 372, 456f., 572, 595, 643
-uhr 529f., 572
Quasar 721f.
-kern 721f.
Quasi|impuls 547, 555
-molekül 218
-neutralität 511, 514, 611
-stationaritätsprinzip 428
Quasiteilchen 476, 502f., 505
-modell 478
quasistellares Objekt 716, 722
Quecksilberdampfgleichrichter 626
Quellung 397
Quellungsvermögen 418
Quencher (Fotochemie) 437f.
q-Raum 479

R

RABI 226
Racemformen 252
Radialverteilungsfunktion 359
Radienquotientenregel 364, 370, 372
Radikal 442, 444
-reaktion 432
radioaktiver Zerfall 652f.
Radioaktivität 29
–, künstliche 32
–, natürliche 29
Radio|astronomie 709
-doppelquelle 720f.
-emission, thermische 710
-frequenzstrahlung 675
-galaxis 718f.
-gramm 167
-graphie 167
-kohlenstoffmethode 168
-metrie 167
-nuklidherstellung 164
-quelle, extragalaktische 729
-quelle, quasistellare 721f.
-reagenzmethode 165
-strahlung 673
-synthese 164
-wellenspektrometer 182
RÄDLER 656
RAMAN 274
Ramanbanden 274
Ramaneffekt 274
–, klassischer 274f.
–, nichtlinearer 275
–, stimulierter 275
Ramanspektroskopie 272
–, Anwendung 276
Ramanstreuung 191
RAMSAY 25, 31
Rand|bedingung, periodische 477, 489
-winkel 412
random walk 85
Raoultsches Gesetz 338
Rationalitätsgesetz 353
Raum, euklidischer 726
–, interstellarer 629
Raum|gitter 79, 352, 356, 478
-gruppe 352, 356
-krümmung 727
-ladung 514, 566, 623
Rauschtemperatur 618
RAYLEIGH 187
Rayleighstreuung 38, 191, 275, 695
READ 513
Reaktanten 419
Reaktion, additive 455
–, asynchrone 440, 444
–, bimolekulare 429, 432f., 434
–, chemische 419
–, chemische (diffusionskontrolliert) 434
–, chemische Klassifizierung 438
–, chemische Stoffwandlung 419
–, elektrochemische 459
–, eutektische 376
–, eutektoide 378
–, fotochemische 437
–, heterogene 455
–, heterolytische 441
–, homolytische 441
–, ionische 441
–, katalytische 436
–, kinetisch kontrollierte 428
–, monomolekulare 430, 432f.
–, pericyclische 440
–, peritektische 377
–, polare 441
–, radikalische 441, 444
–, strahlenchemische 437
–, synchrone 444
–, thermodynamisch kontrollierte 428
–, thermonukleare 647
Reaktions|energie 306
-enthalpie 306

Reaktions|entropie 424
–gleichung 419
Reaktionsgeschwindigkeit, absolute 428f.
– (Lösungsmitteleinfluß) 435
Reaktions|geschwindigkeits-beeinflussung 433
–grad 420
–grenzströme 469
–kinetik 425, 455
–koordinate 431
–ladungszahl 465
–laufzahl 309
–mechanismus 427
–ordnung 425
–schicht 469
–stromstärke 465
–überspannung 467, 469
–volumen 422
Reaktivität (Moleküle) 444
Reaktor, gasgekühlter 179
–kühlung 178
–typ 179
Realfaktor 320
Realstruktur 471
– von Kristallen 509
Rechnen mit Wahrscheinlichkeiten 83
Rechtsschraubenregel 46
Redispergierung 416
Reduktion 439
– der Wellenzahl 479
Reduktionsmittel 463
Redox|elektrode 460
–gleichgewicht 452, 454
–indikator 454
–reaktion 438f.
–titration 454
Reflektanz 590
Reflexion 587, 596
Reflexions|gesetz 588
–klystron 281
–koeffizient 105
–nebel 708, 711
Regenbogenfarben 269
Reggetrajektorien 123
Regolith 670
Reibung, innere 325
Reibungsmodell 664
Reichweite der Kernkräfte 145
REINES 36
Reklameleuchtröhre 626
Rekombination 511, 513, 598, 606f., 660
Rekombinations|koeffizient 513, 606
–kontinuum 608
–strahlung 608
–zentrum 522
Rekristallisation 456
relativistische Dubletts 203
Relavitäts|prinzip 45
–theorie, allgemeine 726
Relaxation 564
Relaxations|effekt 348
–prozeß 189
–verfahren (Reaktionskinetik) 427
Relaxationszeit 189, 348, 380, 517
–, dielektrische 515
Resonanz 36, 58, 80, 175
–, akustische 227
Resonanzabsorption 481
–, paramagnetische 579f.
Resonanz|-Absorptions-Spektroskopie (von γ-Strahlen) 482
–aufspaltung 229
–einfang 175
–integral 238f.
–ramaneffekt 275
Rest|strahleffekt 591
–widerstand 538
RETHERFORD 203
reversibel 304
R-Form 252
Rheologie des Erdinnern 654
Rhombendodekaeder 372
Richardson-Dushman-Gleichung 542
RICHMOND 662
RICHTER 10
Richtungssymbol 354
Riemannsche Geometrie 726
Riesenplanet 677
Riftzone 655
Ringstrom 64
Ringsystem, Saturn 682
–, Uranus 682
RJABININ 637
Robertson-Walker-Metrik 727f.
Rochesche Grenze 704
Rogowskispule 619
RÖMER 37
RÖNTGEN 28
Röntgen|beugung 359, 379
–bremsstrahlung 223
–emissionslinie 223
Röntgenfluoreszenz 224
–spektroskopie 220
Röntgen|hintergrundstrahlung 724
–quelle 706f. 723
Röntgenspektrum 223, 370
–, charakteristisches 223
Röntgen|strahlen 28, 480
–strahllithographie 533
Röntgenstrahlung 707
–, solare 673
Rootspumpe 631
Rotation der Moleküle 256
Rotations|barriere 253
–dauer, siderische 671
–energie 256
–niveaus deformierter Kerne 145
–schwingungsspektrum 262
–spektrum 262
–zustandssumme 313
Roter | Riese 702f.
– Zwerg 705
Rotoxan 256
Rotverschiebung elektromagnetischer Strahlen 482
ROWLAND 182
ROYDS 30
RUBENS 186, 591
Rubin 357
RUBIN 320
Rück|stellvorgang 396
–stoßenergie 481
Ruhenergie 49f., 71
Ruß 379
Russel-Saunders-Kopplung 212, 575
Ruthenocen 248
RUTHERFORD 24f., 30f., 32, 216
Rutherfordmodell 25
Rutil 372
RYDBERG 26
Rydbergkonstante 26, 222
Rydbergniveaus der Moleküle 260
R-Zentrum 384

S

SACKUR 312
Sackur-Tetrode-Gleichung 312
Sagittariusarm 714
Sahaformel 620
Säkulardeterminante 238
säkulare Akzeleration 669
Säkular|gleichung 238
–variation 655
SALAM 133
SALPETER 647
Salz, paramagnetisches 578
–isomerie 251
Sandsackmodell 141, 144
Sandwichverbindung 255
Saphir 357, 529
Satellit 682
–system 682
Sättigungs|dampfdruck 333
–konzentration 337
–polarisation 570

Saturn 682
–monde 683
Satz von der Gleichverteilung der Energie 322
Sauerstoff 285, 367
–, atomarer 657
–, Paramagnetismus 240
–molekül 236
Säure, harte 441
–, mehrwertige 448
–, weiche 441
Säure-Base-|Katalyse 436
–Theorie 436
–Reaktion 439
Säure|dissoziation 421
-dissoziationskonstante 499
–stärke 451
–veresterung 421
SAUTER 554
SAVITCH 34
Schalen, abgeschlossene 210
–modell 143, 203, 217
–quellenmodell 701f.
–struktur 204
Schall, zweiter 349
–absorption 327, 346
–ausbreitung 325, 327, 346
–geschwindigkeit 327, 346
–schwingung 613
–welle 474
Schalt|diode 522
–kreis, integrierter 522, 530
–lichtbogen 627
Schaltungs|entwurf 531
–integration 530
Schaumbrecher 417
Schäume 415, 417
Schäumer 417
Schaumstabilisator 417
Scheibenpopulation 712
Schicht, amorphe 379
–, diffuse 464
–, dünne 472, 517
–, heteroepitaktische 517
–, starre 464
Schichtenaufbau des Erdinneren 652
Schichtsystem, optisches 592
Schmelz|diagramm 376
–druckkurve 375
Schmelzen 366, 639
–, hochpolymere 397
–, inkongruentes 377
–, kongruentes 377
–, unterkühltes 375, 379
Schmelzindex 292
Schmelztemperatur 639
–erniedrigung 377
SCHMIDT 30
Schneeflocken 358
Schockwelle 638, 650
SCHOENFLIES 248
SCHOTTKY 385, 524
Schottky|diode 524
–effekt 523
– -Feldeffekttransistor 529
Schrauben|achse 356
–struktur 299
–versetzung 385
Schraubung 356
SCHRIEFFER 555
SCHRÖDINGER 28, 40, 95, 196
Schrödingergleichung 40, 237, 488
–, radiale 104
–, stationäre 97, 101
–, zeitabhängige 109
Schubmodul 388
Schulz-Flory-Verteilung 286
Schulzverteilung 286
Schutzkolloid 417
Schwankungsquadrat 83f.
Schwarze Löcher 707
SCHWARZENBACH 452
schwarzer|Körper 186, 272
– Strahler 272
Schwarzer Zwerg 705
Schwarzschildradius 707f.
Schwebung 73, 77
Schwefel 285, 367, 375, 455
Schweißen 627
Schweißlichtbogen 627
Schwellenbedingung des Lasers 193
schwerer Wasserstoff 646
schwingende Saite 79f.
Schwingung, elliptische 54
–, erzwungene 58
–, gedämpfte 58
–, harmonische 53
–, lineare 53
– der Moleküle 258
Schwingungs|dauer 53
–mode 80, 473
–quant 91
–quarz 572
–rasse 260
–zustand 473
–zustandssumme 313
Sedimentation 291
Seebeckeffekt 537
Seeligersches Paradoxon 726
SEIGNETTE 569
Seignettesalz 569
Sekundärelektronen 621
–emission 17, 543
–vervielfacher 161, 184, 543, 602
Selbstdiffusion 345
Selbstdiffusionskoeffizient 326
Selen 366f., 603
–gleichrichter 524
Seltene Erden 602
Sensibilisierung, fotochemiche 437
Sensor 507
Sequenz|länge 294
–regel 252
Serien|gesetz 221
–grenze 222
–grenzkontinuum 222
Seyfertgalaxis 716
S-Form 252
SHANNON 82
SHOCKLEY 513
Shubnikovphase 550
Siede|diagramm 339, 376
–punkt 333
–punktserhöhung 339
–temperatur 339
–wasserreaktor 179
–verzug 375
Siegellack 379
Si-Gate-Technik 529
σ-Bindung 239
σ-Orbital 239
Silan 358
Silber|bromid 603
–chlorid 453
– -diammin-Komplex 453
–halogenid 456
–salz 603
– -Silberchlorid-Elektrode 459, 463
Silicat 654
Silicium 357, 362, 368, 384, 386, 495, 506f., 510, 521
–, amorphes 379, 504
–dioxid 372
– -Fotoelement 598
– -Planar-Technologie 528
Sillimanit 374
SIMONS 642
Singularität, kosmologische 725, 727
Sintermagnet 583
Sintern 454
Sinusrille 670
Siriussystem 705
Skalar 46
Skalenfaktor 727
SLATER 369
Slaterdeterminante 108
SLOAN 33
SMEKAL 274
Smith-Purcell-Effekt 273
Smog 442
SMOLUCHOWSKI 13
S_N1-Reaktion 444
SNELLIUS 37
Snelliussches Brechungsgesetz 588
SODDY 16, 30f.

Solar|energie 646
–konstante 671
–zelle 506, 599
Sol 415
–, festes 415
Soliton 486
Solubilisation 407
Solvatation 341, 435
Solvolyse 438
SOMMERFELD 28, 488, 540
Sommerfeldsche Feinstrukturkonstante 197
Sonden|charakteristik 616
–diagnostik 616
Sonne 671, 697
Sonnen|aktivität 658, 673
–atmosphäre 672
–batterie 504, 599
–energie 672
–eruption 675
–fackel 674
Sonnenflecken 673
–zyklus 674
Sonnen|inneres 671
–modell 672
–system 668ff.
Sonnenwind 658, 663, 673
–plasma 685
SOS-Technik 529
SO-Typ 716
SOWOJSKY 227
Spallation 156
Spalt|neutron 174
–produkt 177
Spaltung, spontane 151
Spaltungskammer 160
Spannung, chemische 464
–, eingefrorene 396
–, elektrische 65
Spannungs|-Dehnungs-Diagramm 388
–einheit 65
–koeffizient, thermischer 318
–reihe der Metalle 463
–rißkorrosion 397
–steuerung 528
–verstärkung 526
Speicher|kondensator 571
–ring 114
Spektralanalyse 25
–, qualitative 219
–, quantitative 219
Spektral|klasse 687
–lampe 626
Spektrallinien 687f.
–aufspaltung 225
–profil 618
Spektral|serie 221
–typ 687
Spektro|graph, optischer 617
–heliogram 673
Spektrometer 181
Spektroskopie, zeitauflösende 271
spektroskopischer Wechselsatz 212
Spektrum, elektromagnetisches 69
–, wasserstoffähnliches 222
– eines Moleküls 261
– eines Operators 98
Sperrichtung 520
Sperrschicht 524, 599
–diode 524
–fotoelement 599
–kapazität 527
Sperrstrom 521
Sphäro|lith 396
–kolloid 416
sp^3-Hybridisierung 372
sp-Hybridorbital 241
sp^2-Hybridorbital 241
sp^3-Hybridorbital 241
Spiculen 673
Spiegel, idealer 593
–ebene 247, 354
–kern 134
Spiegelung 247, 354
Spin 52, 64, 100, 117
–abhängigkeit der Kernkräfte 146
Spin-Bahn-Kopplung 52, 142, 147, 202, 496
– im Atomkern 142
Spinecho 228
–methoden 228
Spinell 361, 457, 529
–struktur 654
Spin-Gitter|-Relaxationszeit 227
– -Wechselwirkung 229
Spinhypothese 16
Spin-Spin|-Kopplung, indirekte 230
– -Relaxationszeit 227
Spinresonanz 226f.
–, akustische 231
Spinwelle 584
Spiral|arm 715f.
–nebel 715
–struktur 713f.
Spitzen|diode 521
–transistor 525
Spitzerleitfähigkeit 661
Spreitung 410
Spreitungs|bedingung 410
–druck 411
–geschwindigkeit 411
–koeffizient 410
–spannung 410
–temperatur 549
Spritzgießen 392
Spritzguß 396f.
Spule, stromdurchflossene 66
SQID 559
Stabilisatorröhre 627
Stabilitäts|konstante, Komplex 452
–linie 135
Stadium, radioeruptives 720
Stahl 378
STAHL 10
Standard|atmosphäre 656
–bezugszellspannung 463
–elektrodenpotential 463
–entropie 423
–galvanispannung 460
–potential, chemisches 308
–reaktionsvolumen 309
–wasserstoffelektrode 462
–zustand 306
Stapelfehler 387
STARK 25
Starkeffekt 214, 617
–, linearer 214
–, molekularer 283
– beim Wasserstoffatom 214
STASIW 601
stationärer Zustand 88
Statistik der Kristallelektronen 497
statistische | Physik 471
– Schwankung 81
statistisches Kollektiv 81f.
statistische Theorie der Molwärmen 323
Staub, interstellarer 711
Staudingerindex 292
Steady-State-Theorie 728
STEENBECK 656
Stefan-Boltzmannsches Gesetz 187
Steig|geschwindigkeit in Kapillaren 403
–höhe in Kapillaren 402
Steinsalz 353
Stellungsisomerie 250
Stempel-Zylinder-Kammer 633
Stern, Zustandsgröße 686f.
–atmosphäre 690, 694
–aufbau 697
–dichte 712
–durchmesser 686
–entstehung 700
Sternentwicklung 697, 701
–, Endstadien 705f.
Stern-Gerlach-Versuch 17
Stern|haufen 704
–kontraktion 698
–masse 686
–modell 699
–rotation 689

Stern|schnuppe 685
–schnuppenfall 685
–spektrum 687f.
–system, extragalaktisches 715
Stickoxid 662
Stickstoff 285, 367
–atom, angeregtes 662
–molekül 236
Stilben 252, 438
Stöchiometrie 419
–zahl 419
STOCKBARGER 357
Stoffe, anorganische 268
–, grenzflächenaktive 405, 407
–, hartmagnetische 581
–, kapillaraktive 405, 407
–, kondensierte 327
–, nichtkristalline 356, 370
–, weichmagnetische 581
Stoffwandlung 419ff.
STOKES 271
STONEY 15, 24
Störstellen 384
–beweglichkeit 517
–erschöpfung 512
–kompensation 509
–leitung 506, 508
–reserve 512
Störterm 510
Stoß 47
–, elastischer 47
–, inelastischer 51, 606
–, phasenverändernder 190
Stoß|parameter 52, 216
–querschnitt 323
–theorie 429, 434
–verbreiterung 189f.
–wellen 486, 633, 638, 663, 673, 710
–wellenexperiment 654
–zahl 323
– zweiter Art 218
Strahl, außerordentlicher 595
–, ordentlicher 595
Strahlen|chemie 437f.
–fläche 594
–quelle, radioaktive 167
–schranke 167
–schutz 163
–wirkung 157f., 167
Strahl|geschwindigkeit 594
–richtung 594
Strahlung, charakteristische 28
–, kosmische 706, 712
– im Plasma 608
Strahlungs|bilanz der Hochatmosphäre 658
–dämpfung 189
–detektor 160
Strahlungs|diffusion 609
–druck 70
–einfang 174
–empfänger 274
–feld, interstellares 711
–gesetz 185f.
–gleichgewicht 695
Strahlungsgürtel 666, 668, 681
– der Erde 619
Strahlungs|lebensdauer 189
–leistung 183
–meßgerät 162
Strahlungsquelle, radioaktive 163
Strahlungs|quellen, Infrarotspektroskopie 272
–rekombination 607f.
–temperatur 687
–therapie 168
–transportgleichung 695
Strangeness 119
Strangpressen, hydrostatisches 644
STRANSKI 469
STRASSMANN 34, 154
Stratopause 656
Stratosphäre 656, 659
Streakkamera 184
Streamerkammer 117
Streuprozeß, quantenmechanischer 105
Streuung, elastische 175
–, kohärente 695
–, tief-inelastische 131
–, unelastische (von Lichtquanten) 476
– an einer Netzebene 78
– von Meßwerten 84
Strichgitter 78
Strippingreaktion 155
Strom, anodischer 466
–, elektrischer 63
–, faradayscher 465
–, feldliniengerichteter 666
–, katodischer 466
–, neutraler 130
–, nichtfaradayscher 465
–ausbeute 466
–anteil im Halbleiter 515
–dichte 63, 466, 515
–einheit 63
– -Spannungs-Charakteristik 622
–system 662
–verstärkung 526
Stromverstärkungsfaktor 526
Struktur, ataktische 296
–, isotaktische 296
–, syndiotaktische 296
–brechung 342
– reiner Flüssigkeiten 328
Struktur|theorien des flüssigen Zustands 329
–übergang 639
–untersuchung 480
–verstärkung 342
STUART 369
Stufe, polarographische 468
Stufen|gitter 182
–regel 375
–versetzung 385f.
Sublimation 365
Sublimations|druckkurve 374
–enthalpie 362
Sublimieren 375
Submikrometerbereich 532
Subsolarpunkt 658
Substituenteneffekt 445
Substitution 444
– (chemische Reaktion) 438
–, elektrophile 441
–, gekoppelte 383
–, nukleophile 441
Substitutions|mischkristalle 353, 370, 377, 383
–störstelle 509
Sulfat 374
Super|granulation 672
–haufen 723
Supernovae 693, 702
–ausbruch 712
–überrest 694, 706
Superrotation der Thermosphäre 659
Suprafluidität 350
Supraleiter 488, 641
–anomalie 551
–anwendung 556
– dritter Art 550, 556
– erster Art 550
–kabel 556
–wicklung 556
– zweiter Art 550
Supraleitung 505, 533, 548
Suspension 415, 417
Suszeptibilität 546
– der Paulischen Leitungselektronen 578
–, elektrische 561
–, magnetische 370, 573, 575, 641
Sutherland|gleichung 325
–konstante 325
S-Wellen 652
Symmetrie|element 247, 354
–gruppe 248
–operation 247, 354
–zahl 313
–zentrum 248
Synchronzyklotron 113
Synchrotron 114

Synchrotronstrahlung 706, 717, 719, 722
Synthese, stereospezifische 437
System, abgeschlossenes 303
–, aromatisches 242
–, binäres 375f.
–, disperses 398, 415
–, entartetes 540
–, grobdisperses 415
–, heterogenes 302, 374
–, homogenes 302, 374
–, kolloiddisperses 415
–, kolloides 415
–, unitäres 374
– Blei–Antimon 376
– Eisen–Blei 376
– Eisen–Kohlenstoff 378
– Gold–Silber 377
– Kryolith–Aluminiumoxid 377
Szintillation 600
Szintillations|methode 24
–zähler 161, 602

T

Tafel|beziehung 469
–gleichung 468
–reaktion 469
Tageslichtlampe 626
TAMM 524
TAMMANN 454
Tammanntemperatur 454
Taperpolymer 294
Taschenrechner 529f.
Tautomerie 251
Taxie 296
Technologie der Mikroelektronik 532
TCP-Theorem 132
Teilchen, freies 46
–, hochenergetisches 668
–, neues 123
–, virtuelles 127
–beschleuniger 32, 111ff.
–geschwindigkeit 322
–radius 224
–strom 676
–system 471
–zahl, molare 318
–zahlaustausch 90
Teleskopeffekt 396
TELLER 408
Tellur 367
TELTOW 601
Temperatur 89
–, absolute 318
–,charakteristische (der Rotation) 313
Temperatur, charakteristische (der Schwingung) 313
–, effektive (eines Sterns) 687
–, kritische 320
–abhängigkeit 464
– des Plasmas 610
–koeffizient 538
–verteilung im Erdinnern 653
Tenside 405
–, anionaktive 406
–, kationaktive 406
–, nichtionogene 406
Tensor, metrischer 726
Term|schema 140
–symbol 212
Terrafläche 670
Tetrahedran 256
TETRODE 312
T^3-Gesetz 483
THALES 9
Theorem der übereinstimmenden Zustände 320
Theorie, kinetische 303
– der interionischen Wechselwirkung 343
– der molaren Wärmekapazität 484
– der schwachen und elektromagnetischen Wechselwirkung 133
– des Übergangszustands 467
Thermo|analyse, differentielle 483
–diffusion 136, 326
Thermodynamik 302
–, Aussagegrenzen 424f.
–, chemische 422f.
–, elektrochemische 459
–, phänomenologische 302, 303
–, statistische 303, 310
Thermo|element 184, 274
–gravimetrie 456
–lumineszenz 266
–lyse 442
–plaste 392
–sphäre 656, 658
–topographie 332
Thixotropie 418
THOMAS 208, 637
Thomasfokussierung 113
THOMPSON 238
THOMSON 15f., 24
Thomsonstreuung 695
Thorium 653
Thortveitit 372
Thyratron 627
Thyristor 523
Tiefenbohrung 654
TISZA 349, 569
Titius-Bodesche Reihe 676
Tokamak 648
Tolmanversuch 535
Toluen 382
TONKS 613
Topomerisierung 251
Totalreflexion 75, 170, 589ff., 614
TOWNELEY 10
TOWNES 191
Trägertransportgleichung 516
Trägheits|gesetz 45
–halterung 647, 650
–moment deformierter Kerne 145
Transformationstemperatur 380f.
Transfusion 323
Transistor 525
–bauelement 526
– -Transistor-Logik 531
Translation 247, 356
Translations|symmetrie 471
–vektor, fundamentaler 478
Transmissionskoeffizient 105
Transmittanzen 590
Transporterscheinung in | Gasen 325
– Flüssigkeiten 344
Transport|faktor 525
–prozeß 504
–reaktion 358, 458
–vorgänge 503
Transuran 35
TRAUTZ 429
Treibhauseffekt 678
Trennrohr 326
Tribolumineszenz 266
Tridymit 372, 456f.
Tripelpunkt 375
Tripelpunktsdruck 358
Tritium 646
–methode 168
Trogregion 666
Troktolit 671
Tröpfchenmodell 141
Tropopause 656
T-Tauri-Sterne 701
TTF-TCNQ 536
TTL-Logik 531
TTT-Diagramm 378, 382
Tunneldiode 512, 522
Tunneleffekt 105, 522
–, quantenmechanischer 519
Tunnel|strom 552
–übergang 552
Tunnelung von Cooperpaaren 557
Turbomolekularpumpe 631
Turmalin 596
Tyndall|effekt 415
–streuung 275

U

Überführung 468
Überführungszahl 348
Übergang, erlaubter (Betazerfall) 150
–, optisch erlaubter 221
–, optisch verbotener 221
–, strahlender 514
–, strahlungsloser 514
Übergangs|metalle 547
–wahrscheinlichkeit 110
Überlagerung, inkohärente 77
Überlappung des *s*-Bands 547
Überlappungsintegral 238f.
Über|riese 686
–schallströmung 663
–schußleitung 384
–spannung 466
Überstruktur 396
–phasen 361
UBV-System 688
UHLENBECK 16
Ulichsche Näherung 422
Ultra|filtration 415
–hochvakuum 629, 648
–schallquelle 584
–schallwelle 474
–zentrifuge 291
Umbra 673
Umkehrpunkt 57
Umladung 608
Umsetzung, doppelte 455
Umwandlung, dilative 456
–, displazive 456, 638
–, rekonstruktive 455, 638
Unbestimmtheits|maß 82
–relation 43, 96
Unitäre Symmetrie 124
Universum 724
Unmischbarkeit 376
Unterkühlung 380
Uranium 653
Uranus 682
–mond 683
Ur|element 9
–knall 727
–nebel 685
USANOVIĆ 446
UV-Spektrum 641

V

VAN ARKEL 458
Vakuum|destillation 339
–erzeugung 629
–meßgerät 632
–messung 629
Vakuum|meter, thermoelektrisches 632
–pumpe 629f.
–verfahren 629
Valenz 235
–band 495, 505
–elektron 224
–elektronenzahl pro Gitterbaustein 549
–strich 236
–strichformel 241
–struktur 244, 246
–tautomerie 251
–theorie 235
van-de-Graaff-Generator 111
VAN DER WAALS 244, 319
van-der-Waals-|Bindung 235, 244
– Gitter 361, 364, 370
– Gleichung 637
– Radius 369
van-der-Waalssche Zustandsgleichung 319
van-Hove-Singularität 485
VAN'T HOFF 235
van't-Hoffsche Reaktions|-isobare 309, 422
–isotherme 423
van't-Hoffsches Gesetz der | Gefrierpunktserniedrigung 340
– Siedepunktserhöhung 340
Variationsfeld 655
–verfahren 238
VB-Theorie 242
Vektor 46
–, axialer 46
–potential 554
Velanebelpulsar 702
Venus 678
–wolken 678
Veränderliche (Sterne) 692
Verarmungsbetrieb 528f.
Verbindung, metallorganische 254
Verbindungshalbleiter 384
Verbrennungs|enthalpie 243
–wärme 306, 423
Verbundwerkstoff 472
Verdampferpumpe 631
Verdampfungs|enthalpie 333
–geschwindigkeit 333
Verdunstung 332
Verfärbung, additive 383
–, fotochemische 383
–, subtraktive 383
Verformung, plastische 389
Vergiftung 177
Vergütungsschicht 592, 651
Verhalten, thermodynamisches 14
Verlust|faktor, dielektrischer 394
–winkel 561, 564
Vernetzung, physikalische 392, 394
Verneuilmethode 357
Vernichtungsstrahlung 159
Verschiebung, chemische 230
Versetzung 511
Versetzungs|dichte 386
–ring 386f.
Verstärker, paramagnetischer 523
Verstreckungseffekt 396
Verteilungsfunktion 498, 609
–, radiale 328
Vielkörperproblem der Plasmaphysik 614
Vierelementelehre 9
Vier-Fermionen-Kopplung 129
VILLARD 31
Virgo A 719f.
Virial|gleichung 320
–koeffizient 320
Viskosität 292, 325, 345, 366, 380ff., 633, 654f.
–, dynamische 325
Viskositätskoeffizient 345
VLSI-Schaltkreis 531
Volmerreaktion 469
Voltaspannung 460
Volumen|diffusion 454
–effekt 215
–rekombination 607
Vorwiderstand 622
VSEPR-Regeln 248f.
VUL 569
Vulkanbauten 679
Vulkanisation 392

W

WABER 369
Wagnersche Verluste 562
Wahrscheinlichkeit 81
–, quantenmechanische 94, 98
–, thermodynamische 307, 310
– für Impulsmessung 95
Wahrscheinlichkeits|amplitude 94
–dichte 84
–interpretation der Wellenfunktion 95
WALTON 33, 111
WARBURY 580
Wärme, latente 306
–, spezifische 305

Wärme, spezifische (der Festkörper) 497
Wärmeausdehnung 485
–, Festkörper 486
Wärme|bild 332
–leitfähigkeit 325
–leitung 325, 345, 476, 485, 487, 653
–leitungsmanometer 632
–leitvermögen 487
–leitzahl 325, 487
Wärmekapazität 482
–, molare 482
–, spezifische 482, 611
Wärme|strahlung 186, 653
–stromschalter 554
–theorem von NERNST 307, 310
–tönung einer Kernreaktion 152
–transport in idealen Gasen 487
Wärmewiderstand 487
–, spezifischer 537
WASASTJERNA 369
Washburngleichung 403
Wasser 374, 382
–, Struktur 329
– -Cluster-Ionen 662
– -in-Öl-Emulsion 417
Wasserstoff, metallischer 641
–arme, expandierende 714
–atom 197
–bombe 35, 647
–brückenbindung 246f., 382, 435, 452
–elektrode 450, 460
–konvektionszone 672
–korona 678, 680
–molekülion 238
–spektrum 222
WATSON 300
WAWILOW 596
WEBB 320
Wechsel|licht 274
–maske 532
–stromentladung 625
Wechselwirkung, elektromagnetische 128
–, schwache 129
–, starke 128
–, unelastische 480
–, zwischenmolekulare 245
Wechselwirkungsenergie 56
Weglänge, mittlere freie 87, 323, 517
WEIGERT 701
WEINBERG 133
Weinsäure 252
WEISS 583
Weiße Zwerge 686, 691, 705, 711
Weißsche Bezirke 583
WEKSLER 112, 114
Welt|all, stationäres 728
–modell 725, 727
Wellen, ebene 71
–, elektromagnetische 68, 587, 614
–, harmonische 485
–, magnetohydrodynamische 614
–, stehende 62, 492
–ausbreitung in Kristallen 475
Wellenfunktion 40, 95
–, antisymmetrische 108
–, radiale 104
–, symmetrische 108
– beim Wasserstoffatom 199
Wellen|länge 72
–mechanik 40, 369
–paket 73f., 95
–theorie 37
–zahl 72
–zahlvektor 72, 472
Welle-Teilchen-Dualismus 37, 41, 588
Werkstoffe 471
Wertigkeit 235
Wettergeschehen 656
WIDEROE 33, 112
Widerstand, elektrischer 346, 476, 503, 640
–, spezifischer 536
–, spezifischer elektrischer 488
Widerstands|änderung, magnetische 546
–polarisation 469
–zellen 598
Wiedemann-Franzsches-Gesetz 536, 540
Wiensches Verschiebungsgesetz 43
WILSON 30, 724
Winkelgeschwindigkeit 53
WINKLER, C. 13
Wirkung, katalytische 379
Wirkungs|quantum 42
–querschnitt 87, 152
–radius 368f.
WKB-Tunnelnäherung 543
Wolfram 458
Woodward-Hoffmann-Regeln 440

X

Xenonfluorid 254
Xerogel 418

Y

YANG 150
YOUNG 37f., 194
Young-|Duprésche Gleichung 412
– Laplacesche Gleichung 402
Youngsche Gleichung 412
Youngscher Modul 388
YUKAWA 36

Z

Zählrohr 160
ZEEMAN 25
Zeemanaufspaltung 226, 618
Zeemaneffekt 213, 617, 688
–, anomaler 214
–, molekularer 283
–, normaler 213f.
ZEISE 254
Zelle, ideale galvanische 461
–, reale galvanische 462
Zeit 45
–ableitung quantenmechanischer Mittelwerte 109
–gesetz (Reaktionskinetik) 425
–richtung 45
–standard 195
Zellreaktion 461
ZENER 390
Zener|diode 523
–effekt 518, 523
Zentral|dichte 699
–kraft 57
–temperatur 699
Zentrifugalpotential 59
Zerfall, leptonischer 129
–, nichtleptonischer 129
–, radioaktiver 425f., 652f.
Zerfallsgesetz, Anwendung 157
–, exponentielles 86
Zerfalls|konstante 31, 156
–reihe, radioaktive 31
–wahrscheinlichkeit 140
Zersetzungsreaktion 438, 456
Zerstäuben 379
Zimmdiagramm 290
Zimmermanngleichung 438
Zink|blendestruktur 640
–oxid 603
–sulfid 370, 455, 601
Zinn 365
Zintlphasen 371
Zirconium 458
Zirkulardoppelbrechung 252
Zirkularmaria 670

Zirkulationssystem | der Thermosphäre 659
–, strato-mesosphärisches 659
Zodiakallicht 685
Zone, verbotene 475, 492
Zufallsgröße 83
Zugfestigkeit 634
Zundervorgang 455
Zündspannung 621
Zusammensetzungsvariable 303, 420
Zustand, aktiver 454
–, amorpher 379
–, fotostationärer 438
–, hypotroper kristallinflüssiger 332
–, innerer 303
–, quasistationärer 110
–, stationärer 101
Zustandsänderung 304
–, adiabatische 305, 321
–, endotherme 306
–, exotherme 306
Zustands|begriff, quantenmechanischer 94
–diagramm 60, 357, 374
–dichte 479, 489, 496, 541, 544
–dichte, effektive 507
–dichtefunktion 484, 491, 507
–funktion 303f.
–formen makromolekularer Stoffe 391
Zustandsgleichung 637, 654, 697f.
–, kalorische 305
–, mikroskopische 637
–, phänomenologische 637
–, thermische 304, 332, 637
– der idealen Gase 318
– des Sonnengases 672
– fester Stoffe 488
Zustandsgrößen 303
–, partielle molare 305
–, reduzierte 320
Zustandsgrößen, thermodynamische 312
Zustandssumme 89, 310, 312, 431f.
–, elektronische 313
–, Thermodynamik 430
–, translatorische 312
Zustandssummenberechnung 311
Zustandsvariable 304
Zweielektronenbindung 238
Zweierstoßrekombination 607
Zwei|flüssigkeitsmodelle 349
–komponentensystem 375f.
–niveausystem 90
–teilchenproblem 60
Zwillinge (Baufehler) 388f.
Zwischengitterplatz 603
Zyklotron 33, 112
–frequenz 62
–kreisfrequenz 502
–resonanz 502, 546

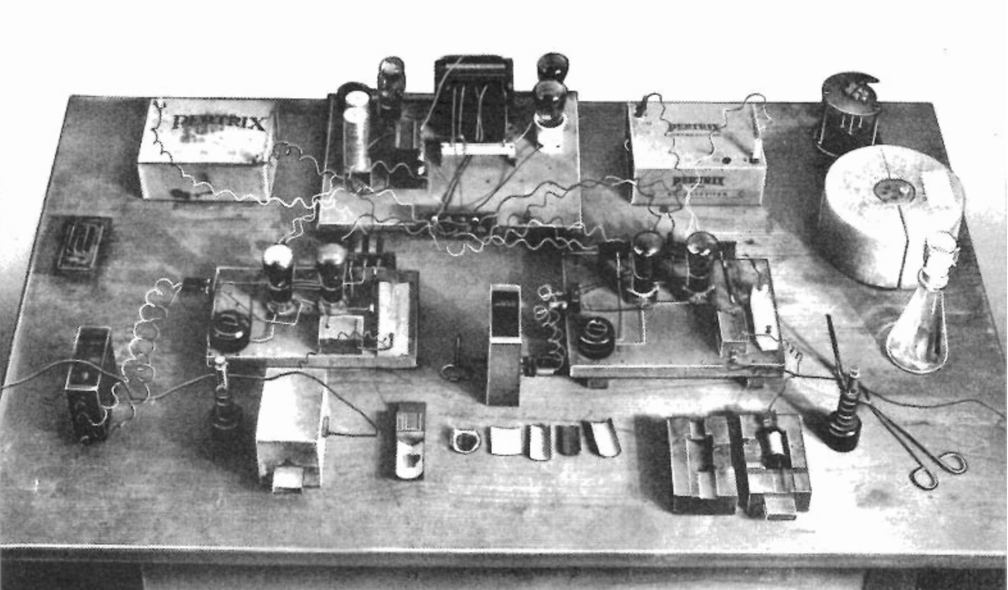


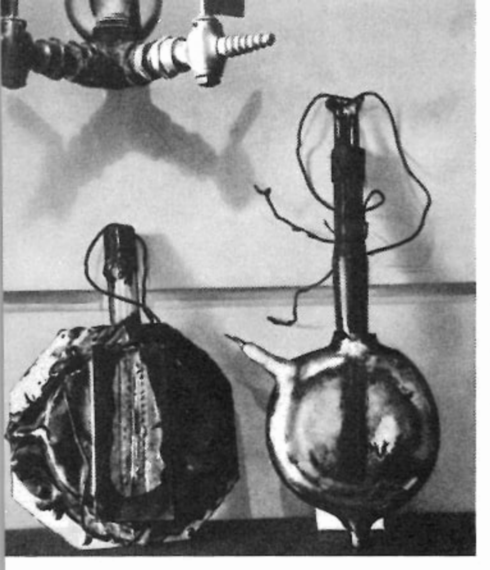
Quæ ſunt in ſuperis, hæc inferioribus inſunt:
Quod monſtrat cœlum, id terra frequenter habet.
Ignis, Agua et fluitans duo ſunt contraria: felix,
Talia ſi jungis: ſit tibi ſcire ſatis!

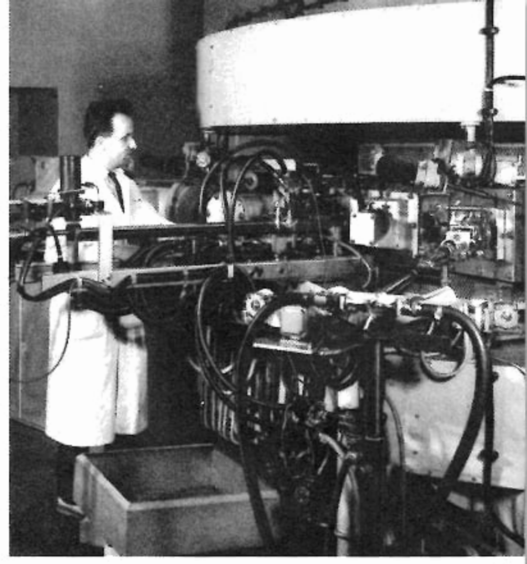
D.M.àC.B.P.L.C.

Tafel 1 Alchemistische Symbolik; Kupferstich 1625 (↑ 1.1.1.2.)

26. – Fév. 96. Sulfate double d'uranyle et de Potasse ...
Papier noir. – Croix de cuivre mince. –
Exposé au soleil le 27. et à la lumière diffuse le 26. –
Développé le 1er mars.

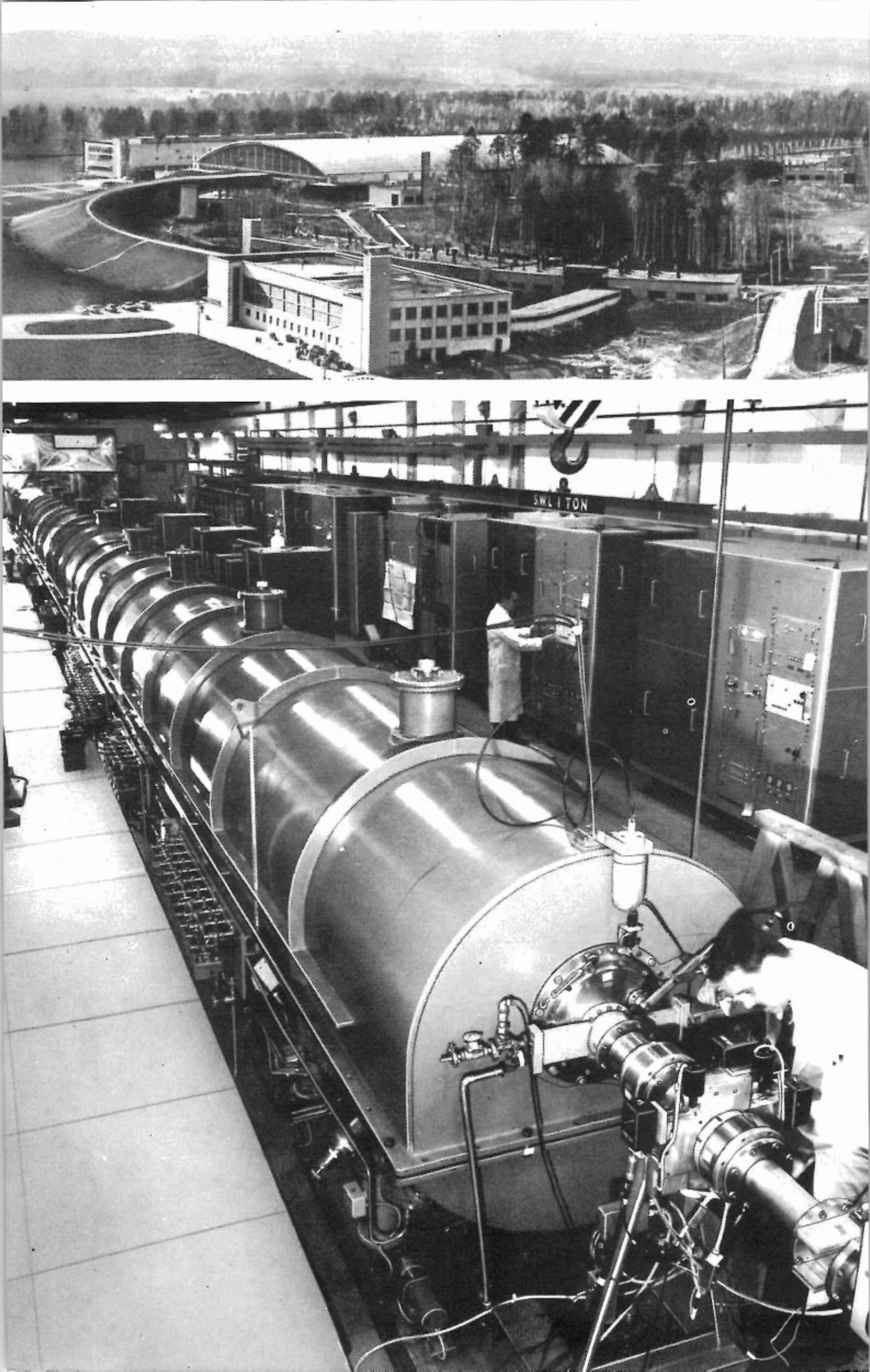






1 | 3 4
2 | 5

Tafeln 2 und 3 **1** Fotografische Platte, die zur Entdeckung der Radioaktivität (↑ 1.1.5.1.) führte, mit Notizen von BECQUEREL – **2** Tisch mit Apparaturen, mit denen HAHN und STRASMANN 1938 die Urankernspaltung (↑ 1.1.5.8.) entdeckten – **3** Die zwei ersten Modelle der Beschleunigungskammer des Zyklotrons (↑ 1.1.5.6.) konstruiert von LAWRENCE und EDLEFSEN – **4** 25-MeV-Zyklotron in Rossendorf – **5** 4-m-Zyklotron zur Beschleunigung schwerer Ionen in Dubna (UdSSR)

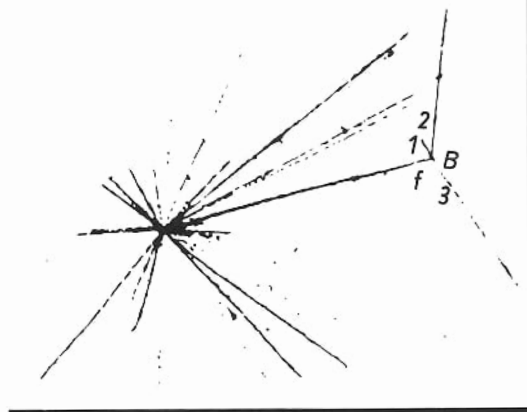

SWL 1 TON

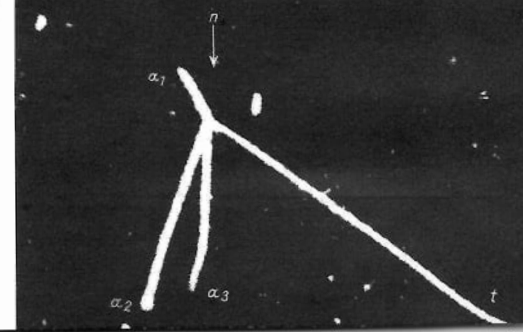


1
4
2 | 3 5

Tafeln 4 und 5 (↑ 1.1.5.6.) **1** 70-GeV-Protonenbeschleuniger im Serpuchower Wald bei Moskau – **2** 50-MeV-Protonenbeschleuniger vom Alvarez-Typ der CERN – **3** Teilansicht des Beschleunigungsrings mit Ablenk- und Fokussierungsmagneten beim Serpuchower Protonenbeschleuniger – **4** Erste Beobachtung eines Hyperfragments mit Hilfe einer Kernemulsionsaufnahme; das Kernbruchstück *f* zerfällt im Punkt *B*, die Messung an den Spuren *1*, *2*, *3* weist daraufhin, daß das Λ°–Hyperon eine Komponente des Kernbruchstücks war – **5** Zertrümmerung eines Stickstoffkerns durch ein Neutron n mit einer Energie von 15,7 MeV in drei α-Teilchen und ein Triton *t*



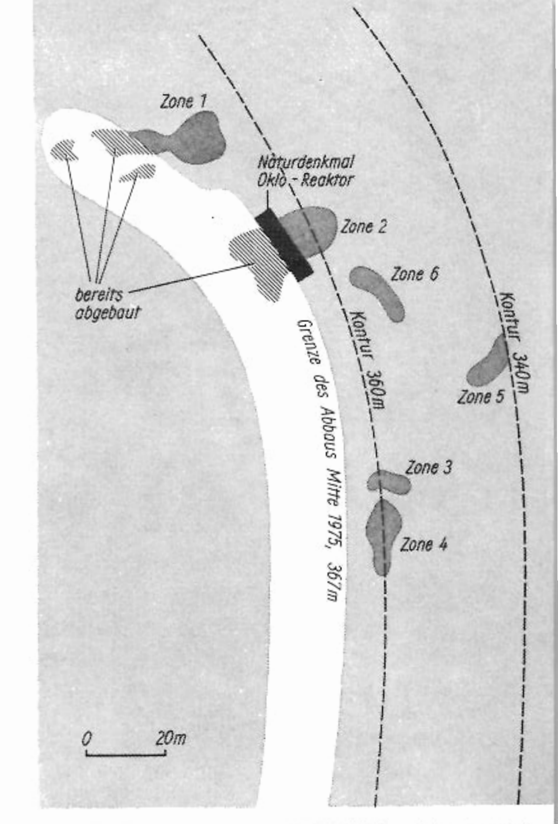


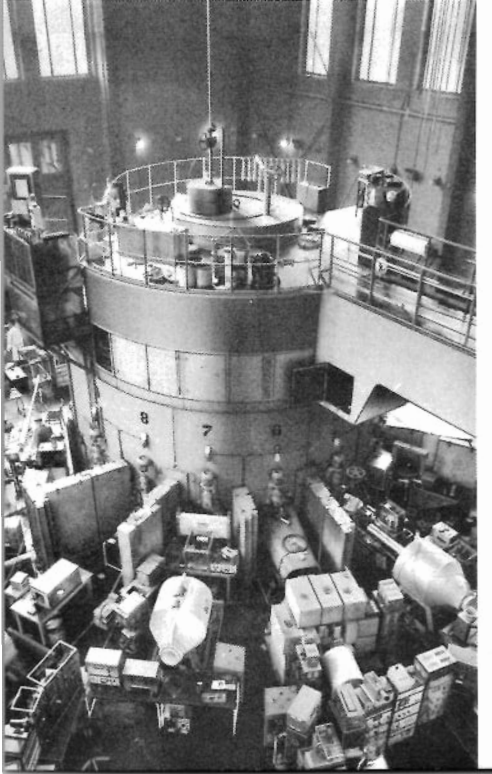


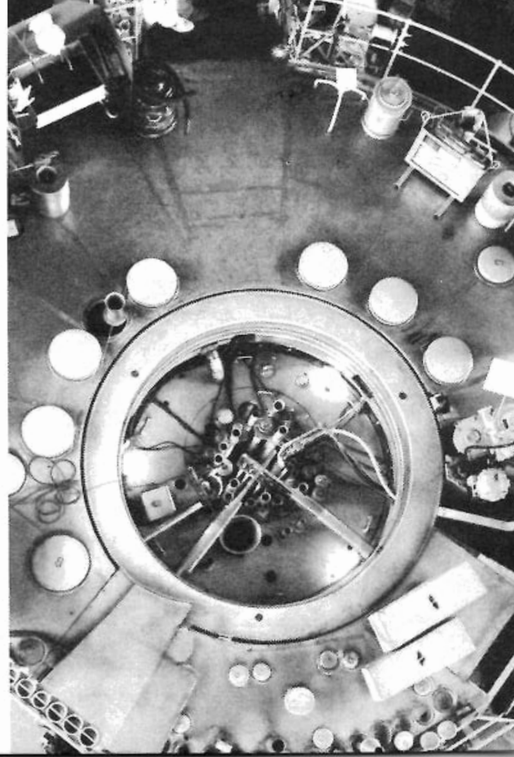


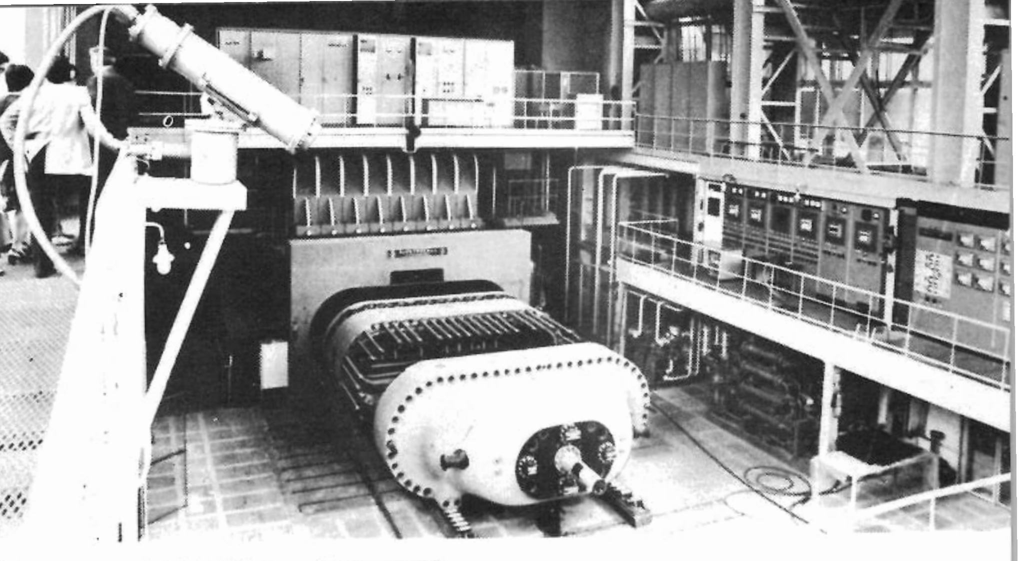
1 | 2
3 4

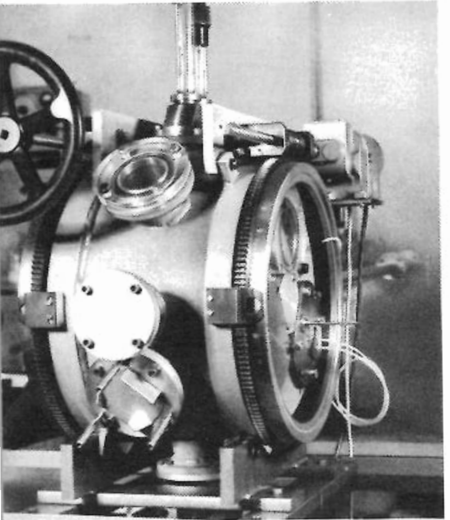
Tafeln 6 und 7 (↑ 1.1.5.8.) **1** Natürlicher Kernreaktor in Oklo (Republik Gabun); durch das Zusammentreffen mehrerer Zufälle konnten in sechs Reaktorzonen des Uraniumerzes natürliche Kernreaktoren entstehen und mindestens 150000 Jahre arbeiten, die Moderation erfolgte durch Wasser und die abnehmende U-235-Konzentration; ein Teil der Reaktorzone *2* liegt im unteren Teil des links im Vordergrund sichtbaren Hangs (vgl. mit **2**) – **2** Draufsicht auf den nördlichen Teil des Oklo-Tagebaus (Zustand 1975); die Reaktorzonen *1* und *2* wurden vor der Entdeckung z. T. abgebaut, die Zonen *3* bis *6* liegen etwa 10 m unterhalb der derzeitigen Sohle – **3** Forschungsreaktor in Rossendorf – **4** Blick in den Reaktorkern des Forschungsreaktors





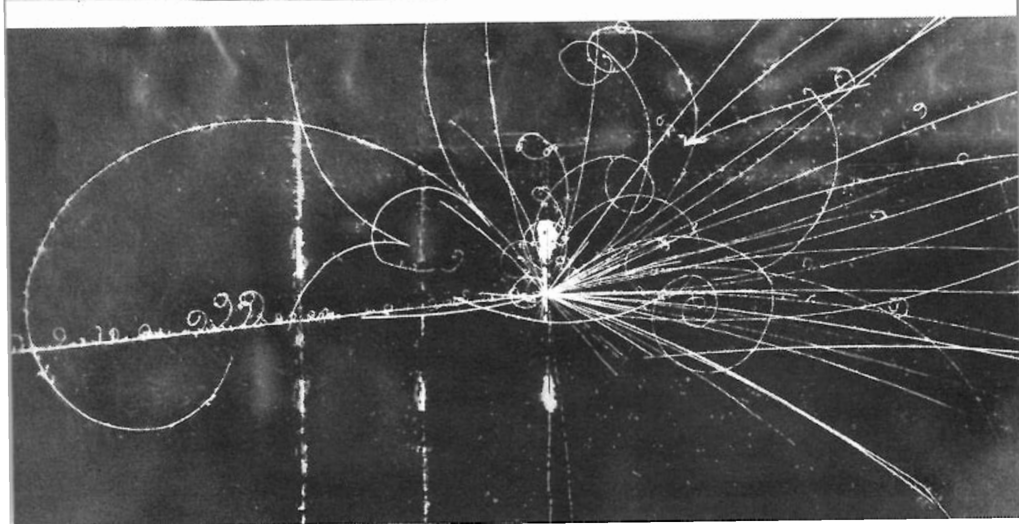






1	4
2	
3	5

Tafeln 8 und 9 (↑ 3.2.1.2.2.) **1** Blasenkammer SKAT in Protwino bei Serpuchow (UdSSR) – **2** Kleine Blasenkammer in Rossendorf – **3** Blasenkammeraufnahme; der aus C-12-Kernen bestehende Strahl (50 GeV) durchdringt drei hintereinanderstehende Wolframfolien; die entstehenden positiven Teilchen werden nach rechts, die negativen nach links abgelenkt – **4** Blick auf die 2-m-Wasserstoff-Blasenkammer von CERN – **5** Blasenkammeraufnahme der Wechselwirkung eines negativen K^--Mesons mit einem Proton; es entsteht ein Ω^--Hyperon, das kurz darauf zerfällt; daneben Reaktionsgleichung und Rekonstruktion der Blasenspuren (↑ 1.1.6.2.).



$$K^- + p \rightarrow \Omega^- + K^0 + K^0 + \pi^+$$

$K^0 \rightarrow \pi^+ + \pi^-$

$\Omega^- \rightarrow \Lambda^0 + K^-$

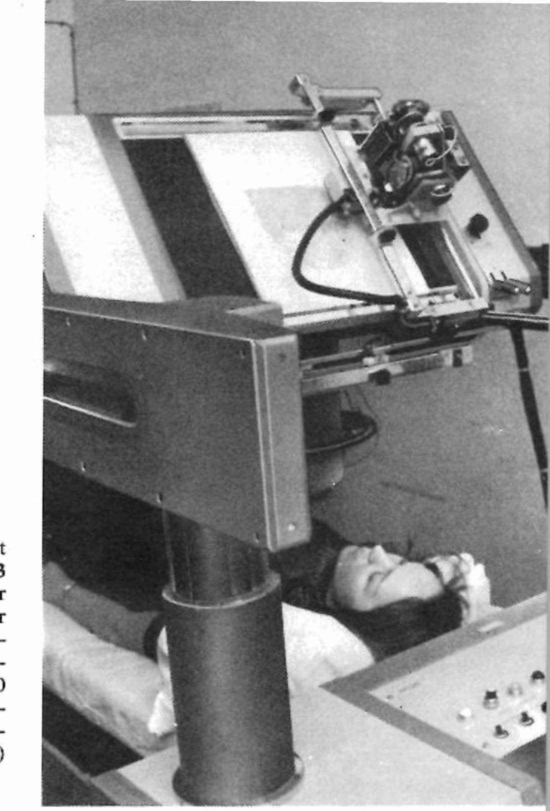
$K^- \rightarrow \pi^- + \pi^0$

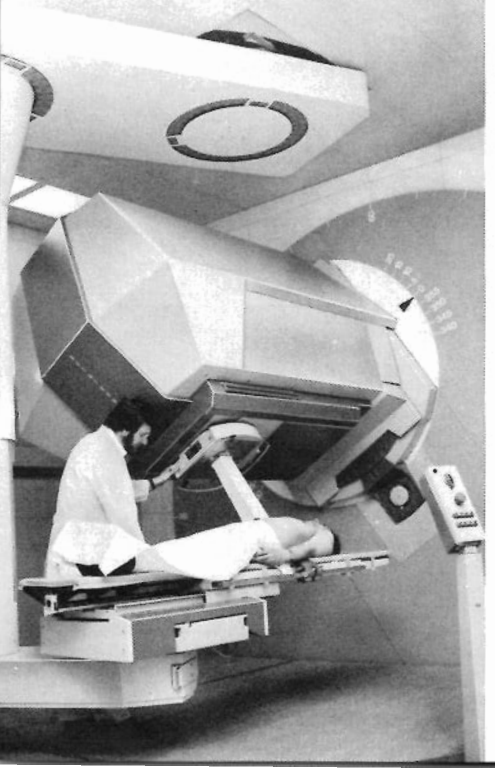
$\Lambda^0 \rightarrow p + \pi^-$


TB 2
1/3

1 | 4
2 3 | 5 6

Tafeln 10 und 11 (↑ 1.1.5., 3.4.3.) **1** Arbeiten mit radioaktiven Stoffen in einer heißen Zelle – **2, 3** Durch Autoradiographie sichtbar gemachter Stofftransport in Pflanzen; radioaktiver Phosphor gelangt über die Stengel zu den Blatträndern – **4** Scintiscanner zur Diagnose krankhafter Prozesse innerer Organe (↑ Tafel 66) – **5** Betatron 500 zur Bestrahlung mit Röntgen- und Elektronenstrahlen (↑ 3.1.2.) – **6** Ganzkörper-Computer-Tomograph zur Röntgendiagnostik (↑ Tafel 66)







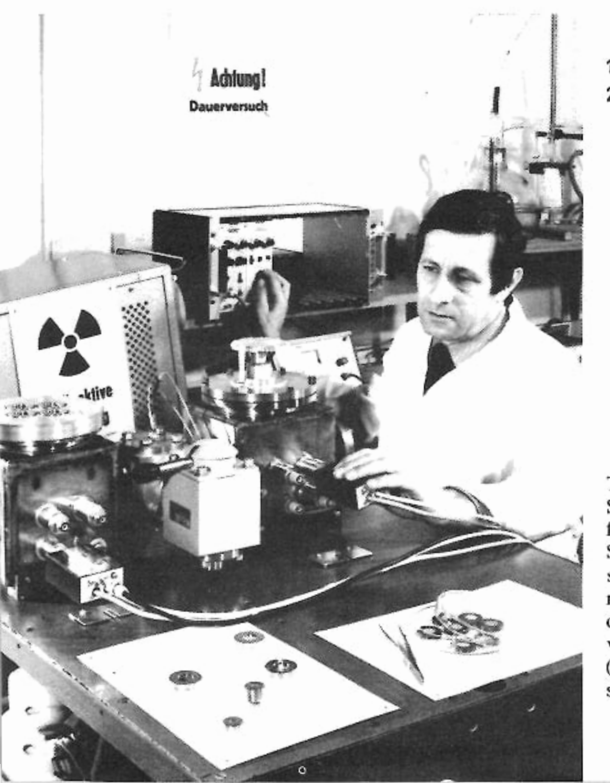
β-, γ- Zählrohre

γ- und Höhenstrahlungszählrohre

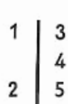
Fensterzählrohre (α-, β- und Röntgenstrahlung)

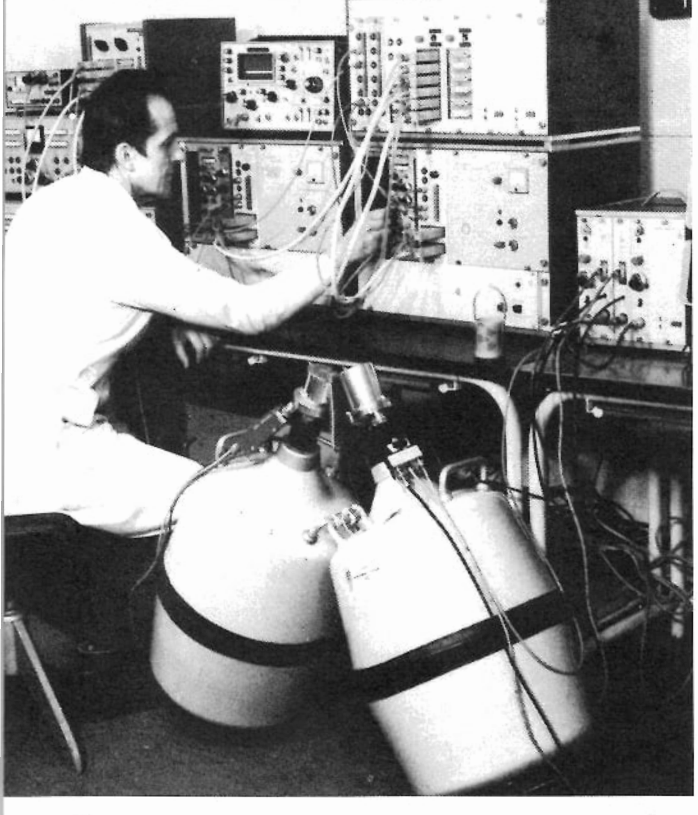
Flüssigkeitszählrohre

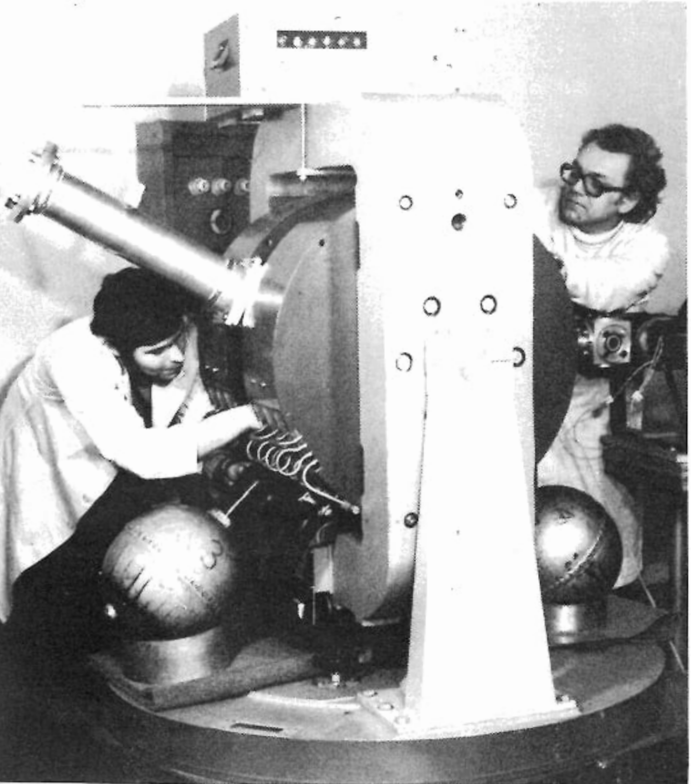
1 | 3
2 | 4 5



Tafeln 12 und 13 **1** Zählrohre als Strahlungsdetektoren – **2** Prüfstand für Halbleiterdetektoren (↑ 8.3.) – **3** Szintillationsteleskop im Baksan-Observatorium (UdSSR) zum Neutrinonachweis – **4** Blasenkammeraufnahme des seltenen Prozesses einer Neutrinowechselwirkung mit einem Elektron (↑ 3.2.2.1.) – **5** Skizze zu der nebenstehenden Blasenkammeraufnahme







Tafeln 14 und 15 (↑ 3.2.1.2.2.)
1 γ-Meßplatz mit zwei Germaniumdetektoren zur Aufnahme von $\gamma\gamma$-Koinzidenzspektren (KfK Rossendorf) – **2** Anordnung zur Messung magnetischer Momente am Teilchenstrahl des Rossendorfer Zyklotrons U-120 – **3** Spektrum der γ-Strahlung, die beim Beschuß von Se-78 mit α-Teilchen einer Energie von 27 MeV am Rossendorfer Zyklotron emittiert wird (Messung mit einem 50 cm³ großen Germaniumdetektor); die Energien der γ-Linien, die Übergängen in Kr-80 entsprechen, sind in keV angegeben (dargestellter Ausschnitt mit höchster Auflösung) – **4** Spektrometer für die Wechselwirkung relativistischer Kerne (d, α, C-12) mit Targetkernen aller Art; *rechts* vom Magneten bestrahltes Target, *links* Fokussierungslinsen und Detektor – **5** Čerenkov-Massenspektrometer „Photon" in Verbindung mit dem Synchrophasotron in Dubna zur Untersuchung der Wechselwirkung von negativen Pionen und Protonen

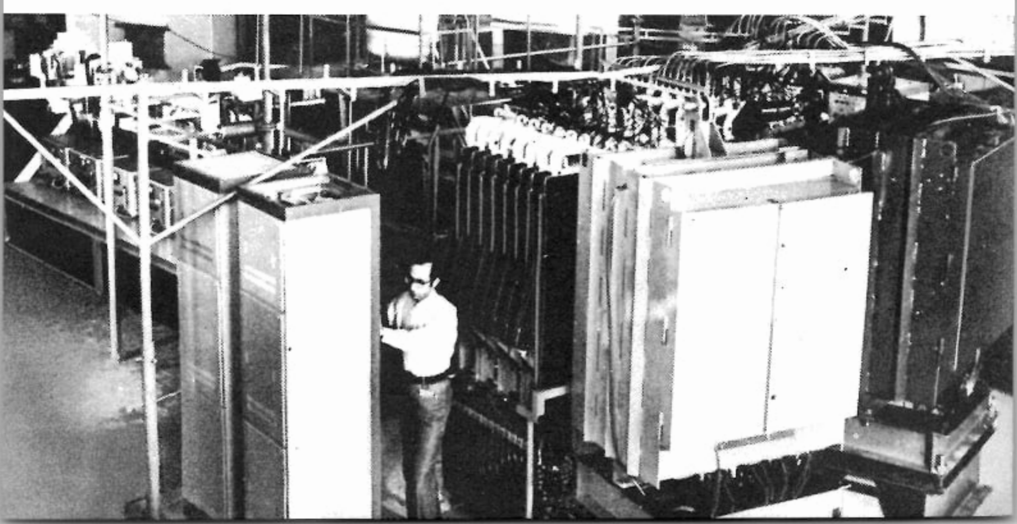
Se-78 + 27 MeV α-Teilchen θ = 125°

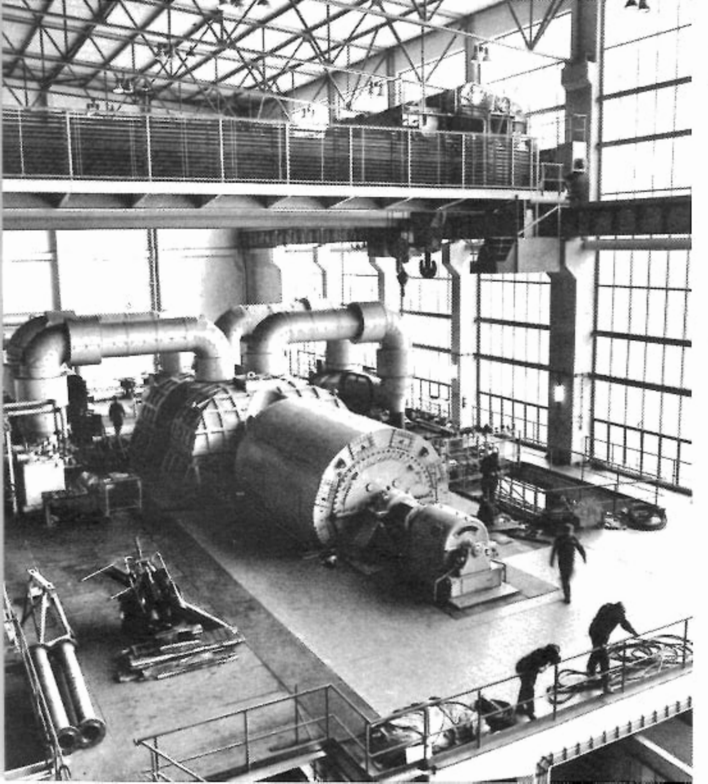
$\cdot 10^4$ Ereignisse pro Kanal

Kanalzahl

Kr-79, Se-77, Br-80, 182,3, Br-79, Br-80, 248,6, 271,0 + Br-80, 289,8, 313,3, 351,8, 382,2, 420,3, Kr-78, 488,4, β⁺, 516,8, 531,7, 539,8, 553,6, 582,7, 596,0, 616,6, 628,5, 639,6, 647,2, 686,0, 694,8, 698,7, 709,8, 722,1, Kr-82, 780,3, 819,4, 834,7, 862,9, 871,4, 889,4, 893,3, 948, 955,8, 964,4, 967,6, 975,5, 980,4, 1004,9, 1017,7, 1032,5, 1043, Kr-82, 1059,5, 1101,5, Kr-82, 1121,5, 1138,2, 1166,5, 1171,4, 1189,6, 1223,3, 1238,2, 1256,2, 1308,1, 1332,0, 1423,4

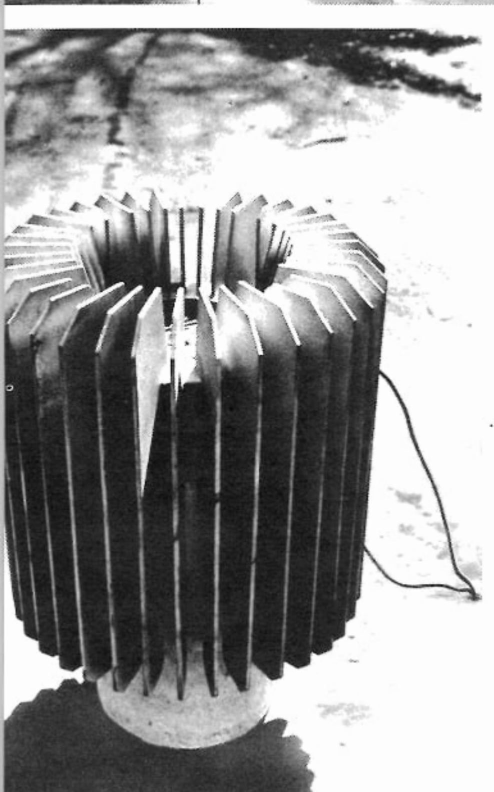
486,0, 488,4, 490,5

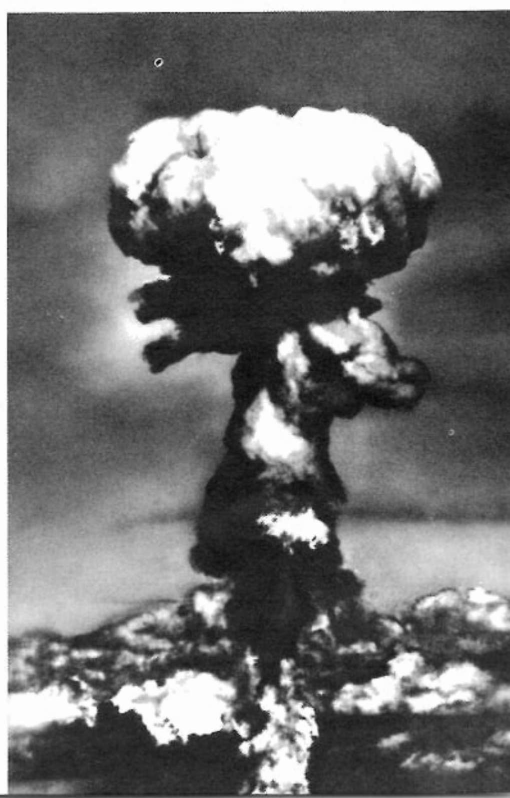


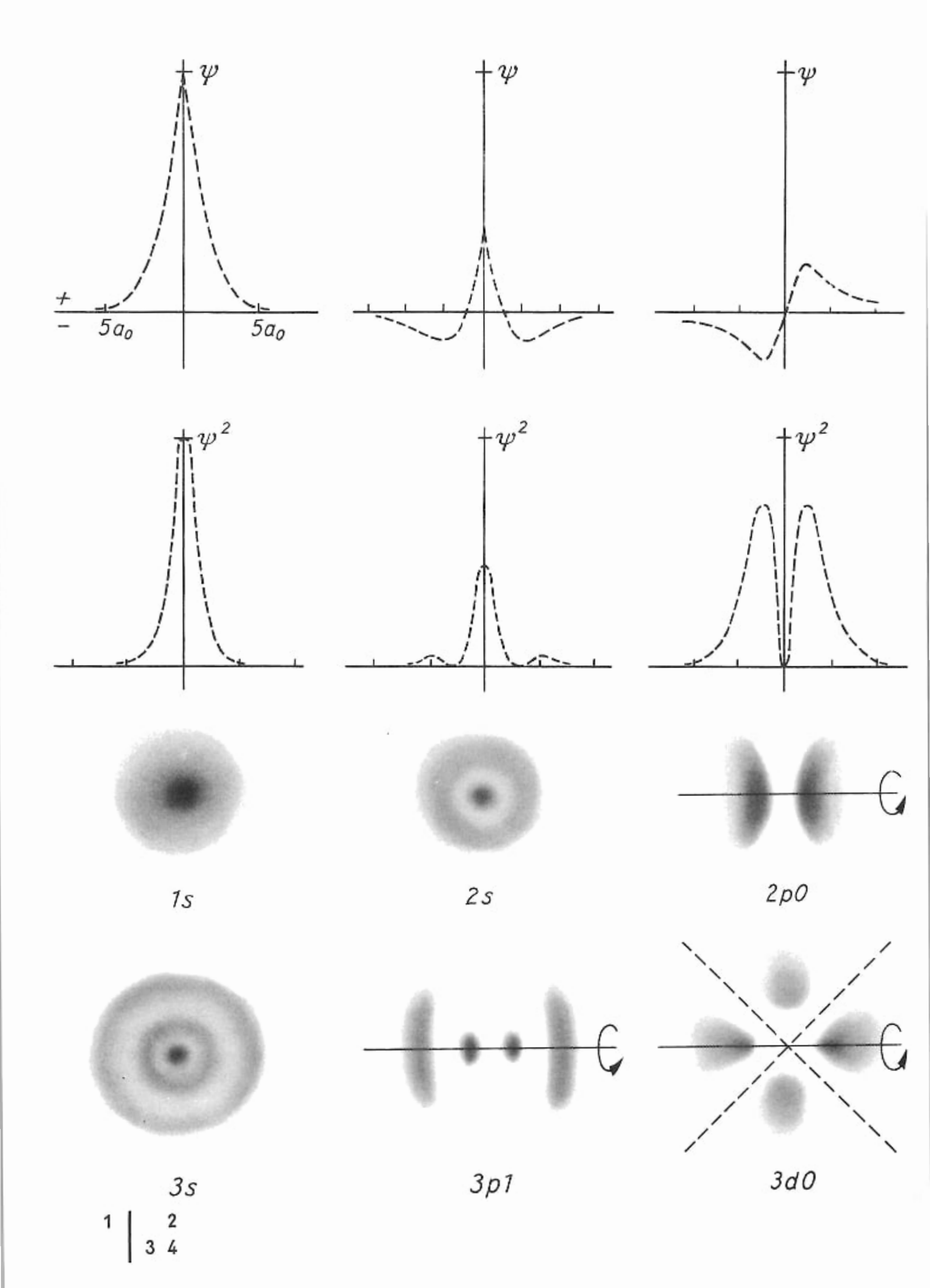


1 | 3
2 | 4 5

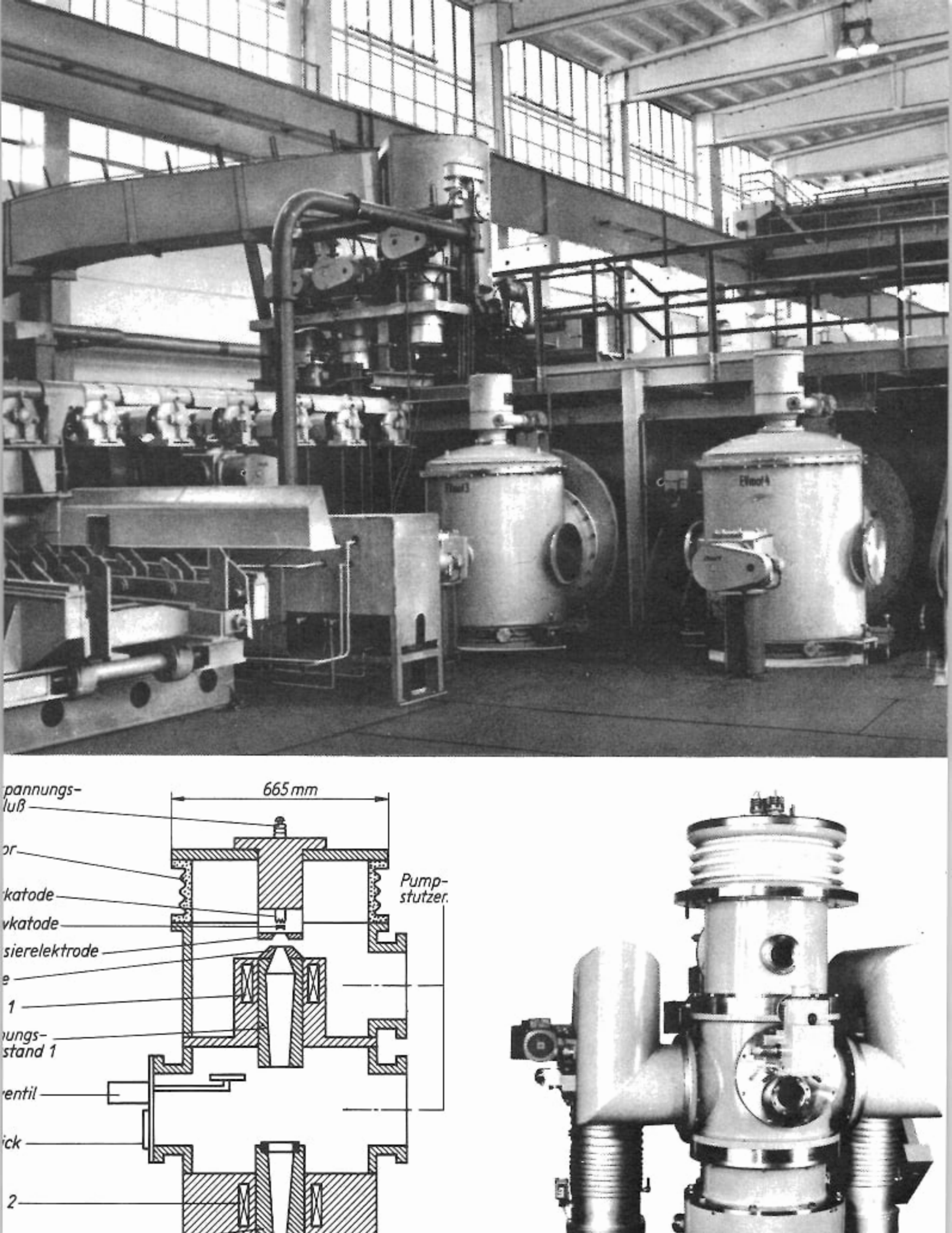
Tafeln 16 und 17 **1** Französisches Kernkraftwerk bei Chinon – **2** Turbine und Generator im Kernkraftwerk bei Rheinsberg († 3.5.) – **3** Isotopentrennanlage († 3.3.1.4.) in Oak Ridge (USA) – **4** In der UdSSR entwickelte Isotopenstromquelle Beta-2, die jahrelang wartungsfrei Strom liefert – **5** Atompilz nach einer Kernbombenexplosion († 1.1.5.9.)



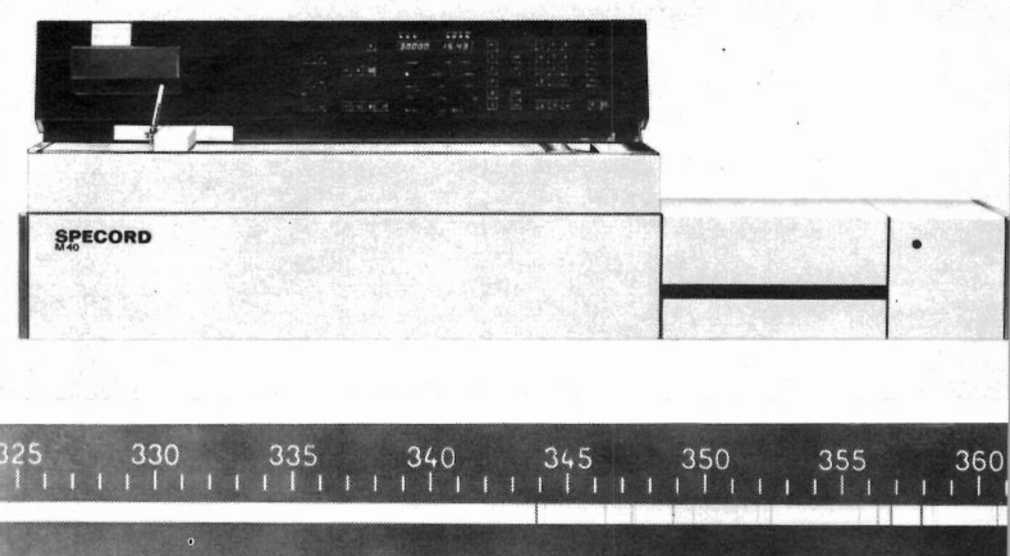


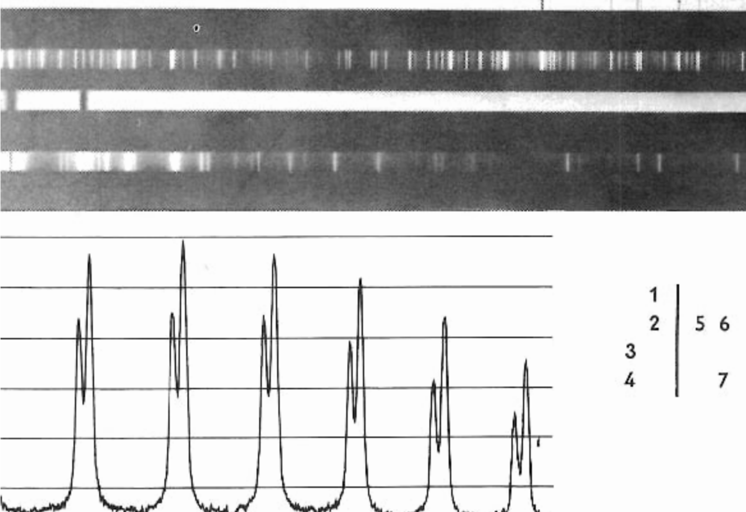


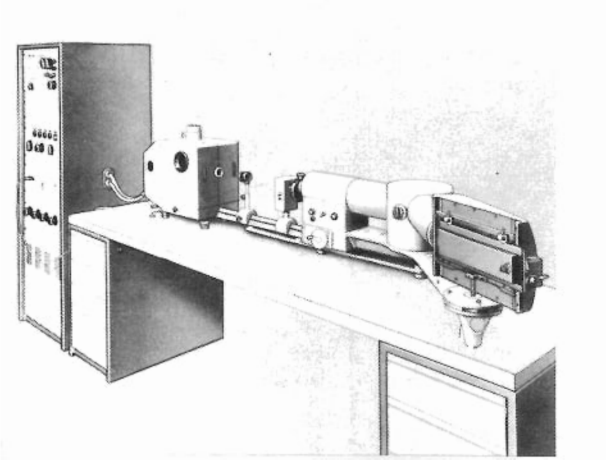
Tafeln 18 (↑ 4.2.2.4.) **und 19** (↑ 9.3.) **1** Wellenfunktionen und Wahrscheinlichkeitsverteilungen beim Wasserstoffatom; Werte für ψ, ψ^2 und räumliche Verteilung für die Zustände 1s, 2s und 2p0; räumliche Verteilung für die Zustände 3s, 3p1 und 3d0; für 2p0 geben ψ und ψ^2 den Verlauf längs der Rotationsachse an, die durch einen *gekrümmten Pfeil* gekennzeichnet ist – **2** Elektronen-



strahlanlage zum Vakuumschmelzen – **3** Schema einer Hochleistungselektronenkanone – **4** Hochleistungselektronenkanone vom Typ EH 1200/50 mit 1200 kW Strahlleistung bei 50 kV Beschleunigungsspannung (entwickelt im Forschungsinstitut M. v. Ardenne, Dresden)

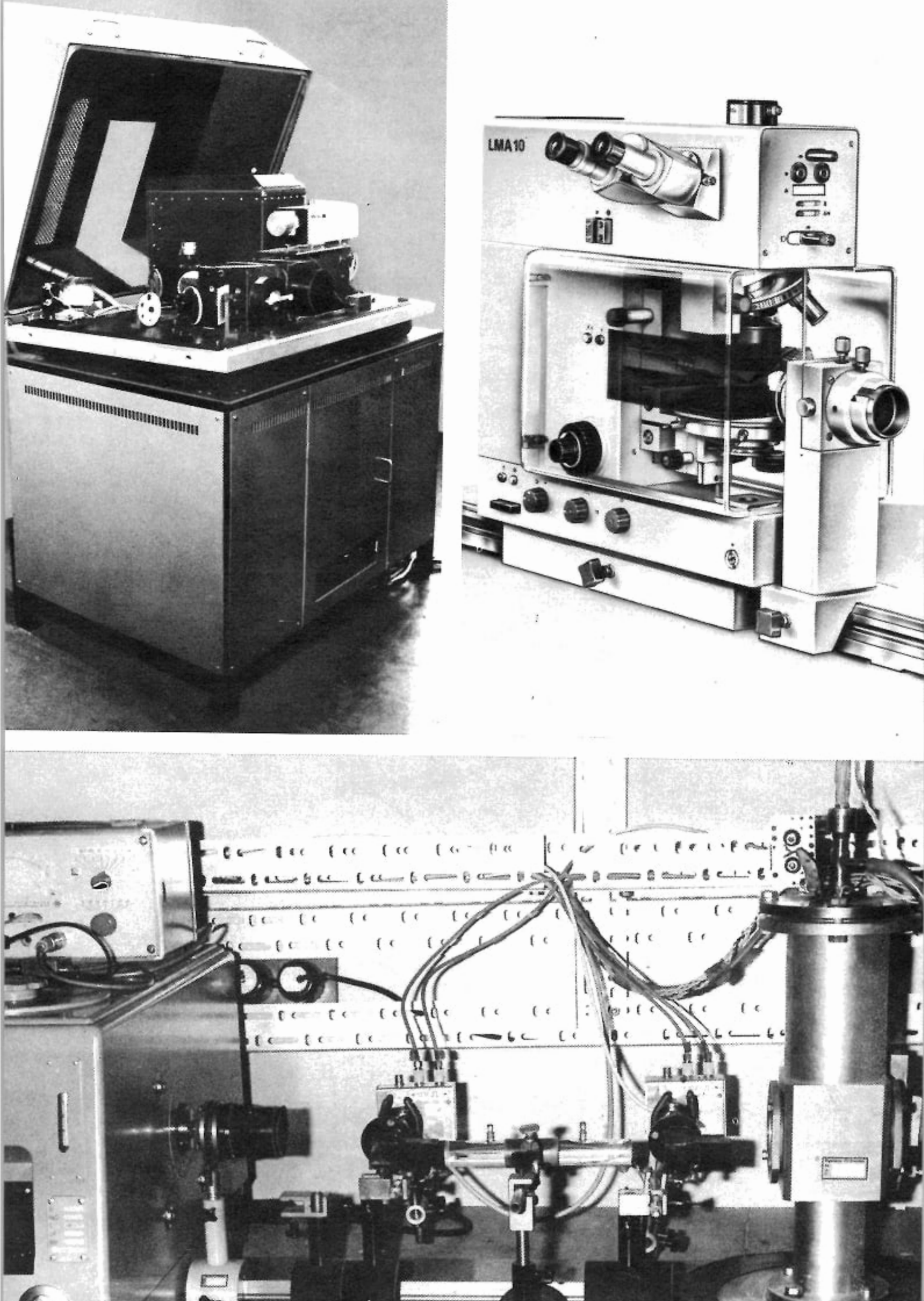






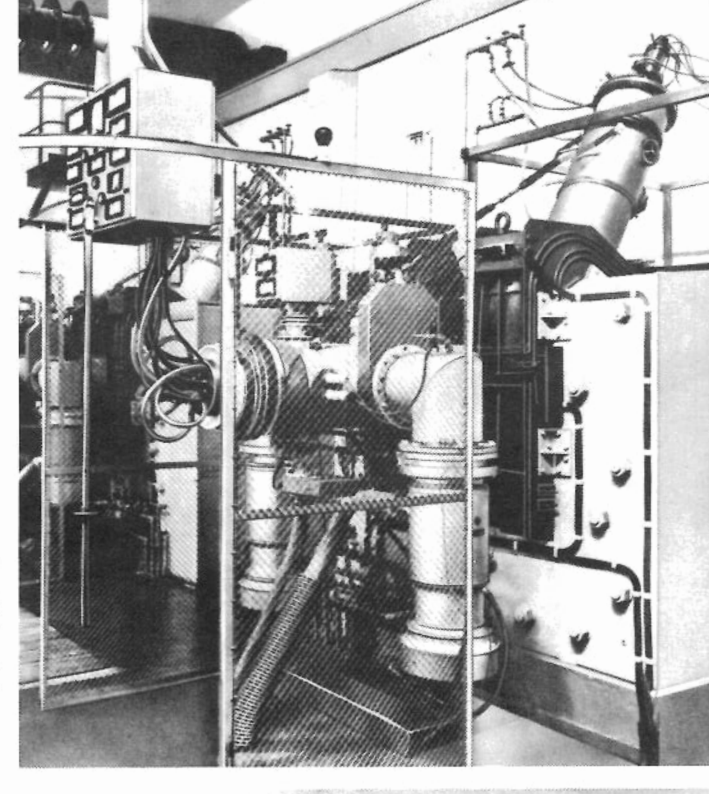
	1	
	2	5 6
3		
4		7

Tafeln 20 und 21 (↑ 4.1.4.7.5., und 4.3.2.1.1.) **1** UV-Spektralphotometer Specord M 40 – **2** Mit einem Lasermikrostrahlanalysator aufgenommene Emissions- und Absorptionsspektren von niedrig legiertem Stahl und Elektrolytkupfer; *oben* Absorptions- und Emissionsspektrum von Stahl, *unten* Absorptions- und Emissionsspektrum von Kupfer – **3** Infrarotspektrum von Chlorwasserstoff – **4** UV-Spektrograph Q 24 – **5** Fourier-Transformations-Infrarotspektrometer – **6** Laserstrahlmikroanalysator LMA 10 – **7** Infrarotabsorptionsmessung mit Laserdiode als Strahlungsquelle


LMA 10

$^{40}Ar^{++}$
^{20}Ne
^{19}FH
$^{18}OH_2$
$^{16}OD_2$
$^{15}NHD_2$
$^{14}ND_3$
$^{13}CD_3$
$^{12}CD_4$

1 | 3
2 | 4



Tafeln 22 und 23 (↑ 1.1.3.5.)
1 Massenspektrometer mit Spektrum – **2** Elektronenresonanzspektrometer – **3** Magnetischer Massentrenner – **4** Experiment mit einem Helium-Neon-Laser

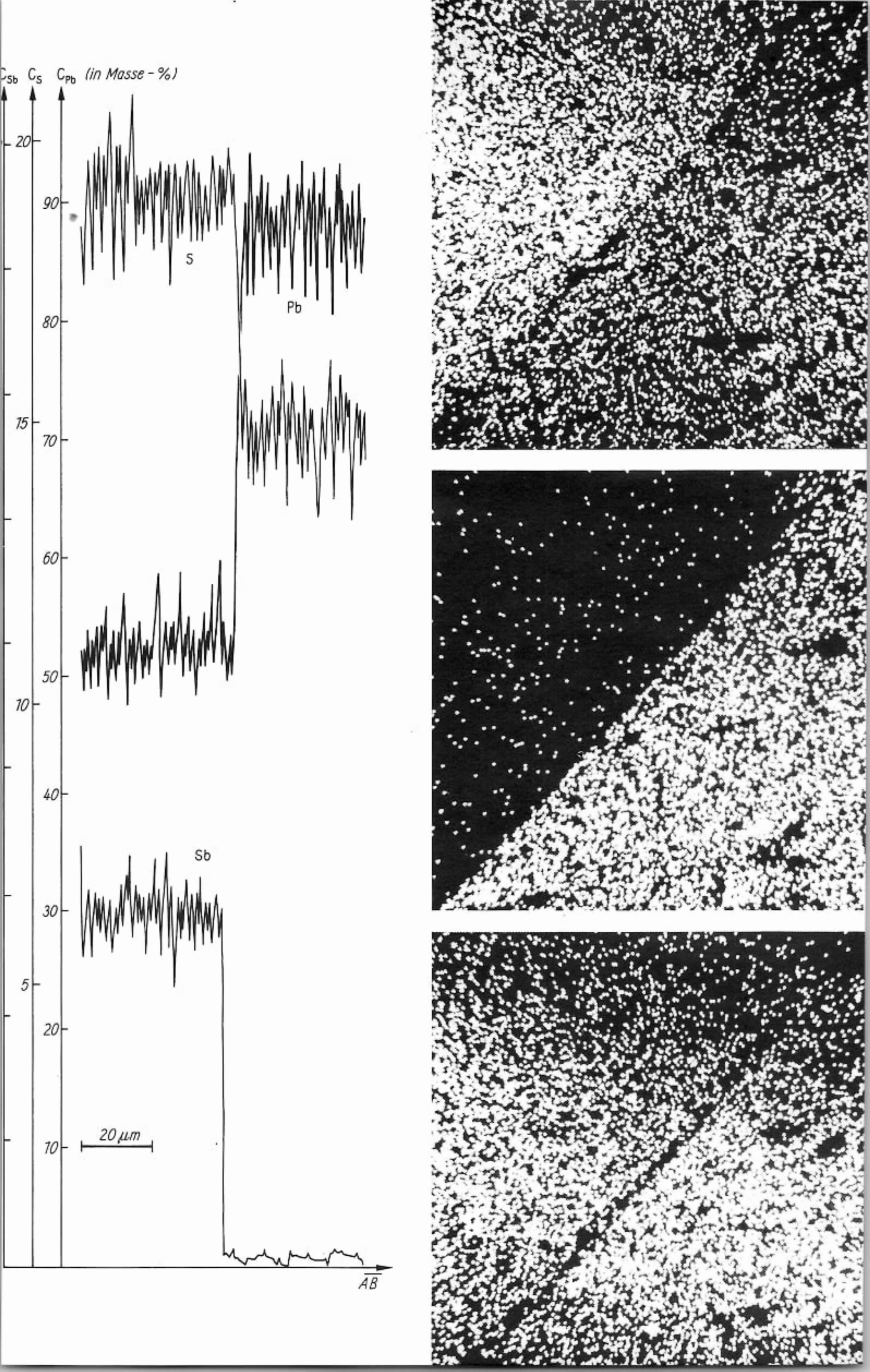
1 | 3
4
2 | 6 5

$\overline{AB}$

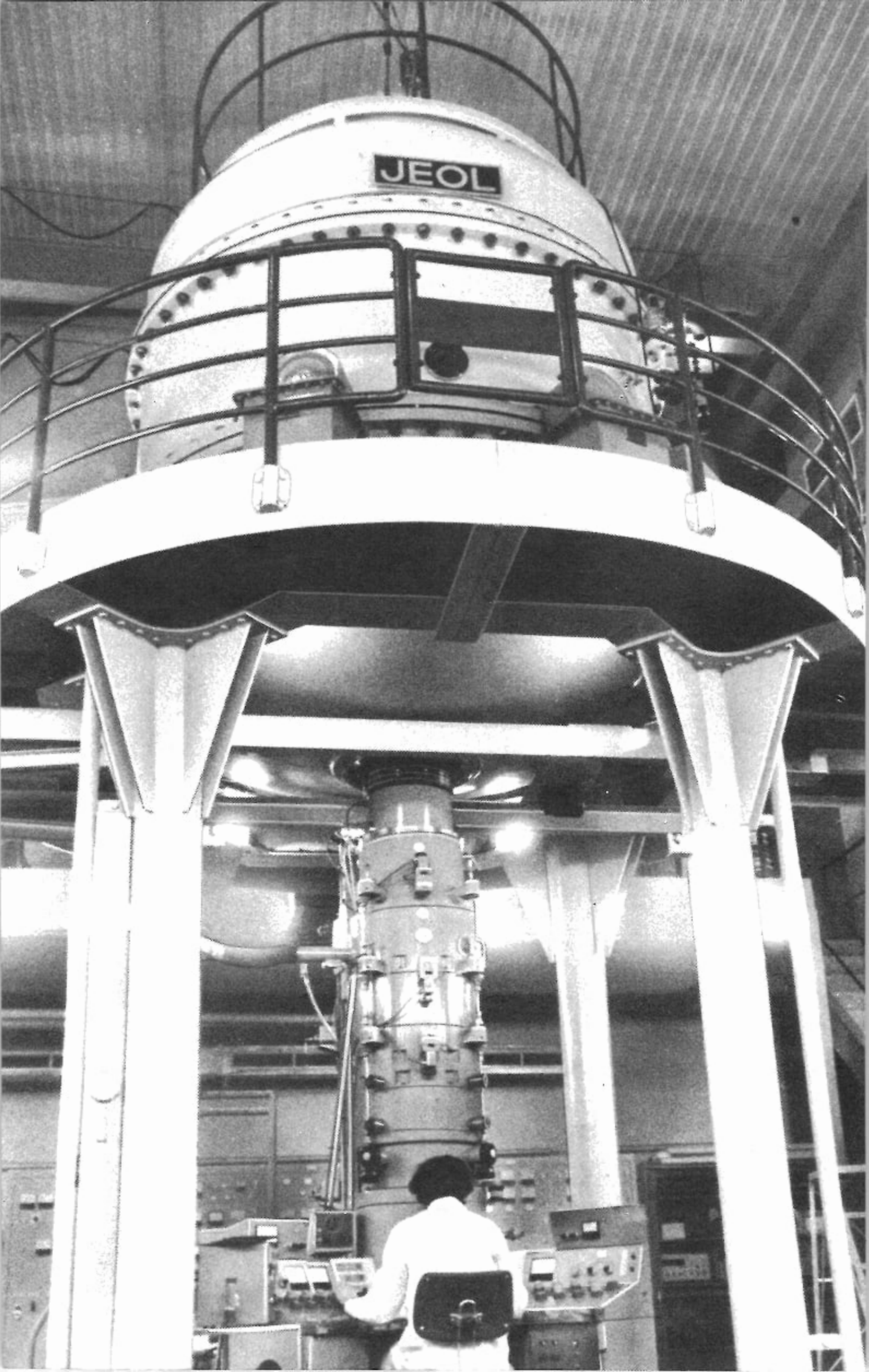
50 μm

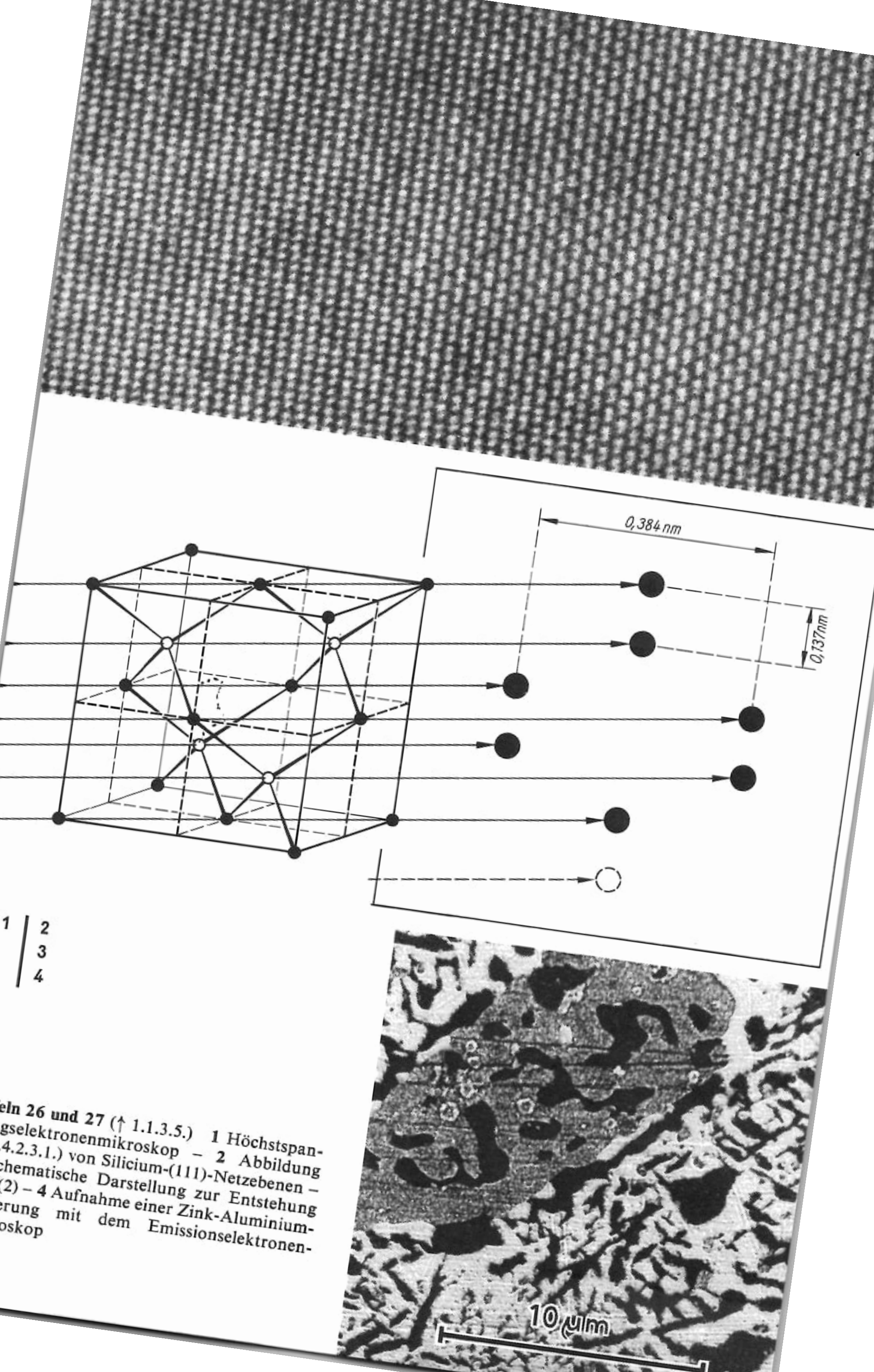


Tafeln 24 und 25 (↑ 1.1.3.5.) **1** Röntgenstrahlmikroanalysator – **2** bis **5** Mit einem Mikrostrahlanalysator erhaltene Rasterbilder einer Erzprobe bestehend aus Bleiglanz (PbS) und Semseyit ($Pb_9Sb_8S_{21}$); Rasterbild der absorbierten Elektronen (**2**), der Pb-M_α-Strahlung (**3**), der Sb-L_α-Strahlung-(**4**) und der S-K_α-Strahlung (**5**) – **6** Massekonzentration c von Pb, Sb und S längs der Linie AB der Probe (**2**)

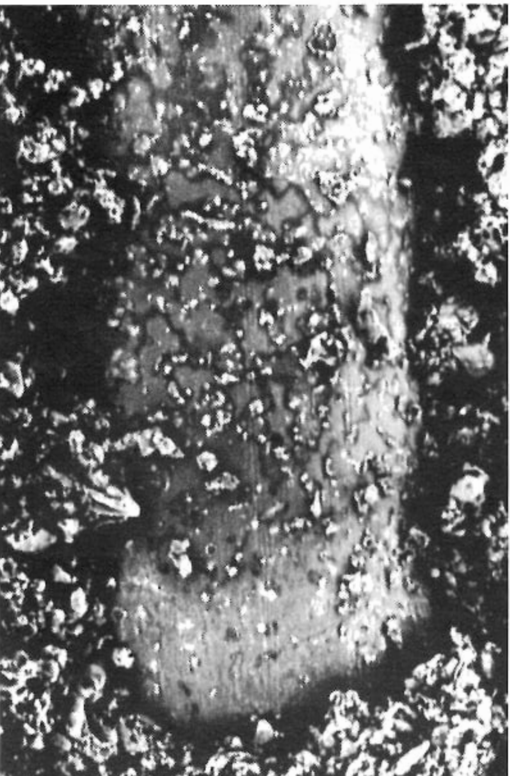


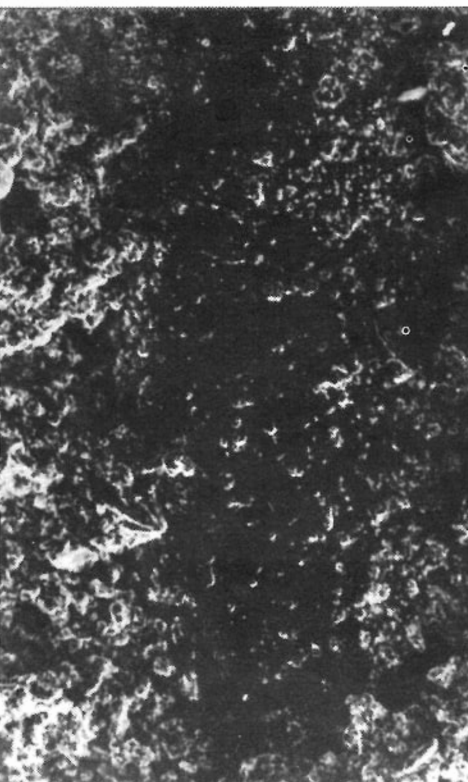
C_{Sb} C_S C_{Pb} (in Masse - %)
20
90
S
Pb
80
15
70
60
50
10
40
Sb
30
5
20
20 µm
10
AB


JEOL



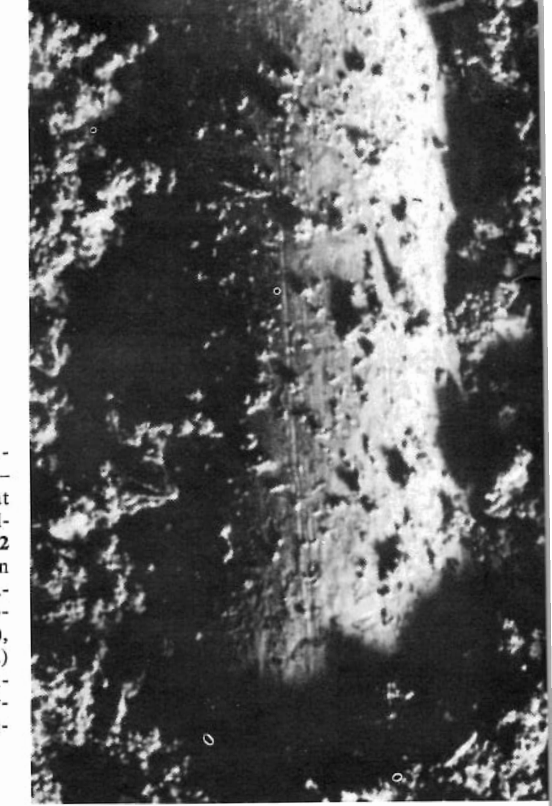
eln 26 und 27 (↑ 1.1.3.5.) **1** Höchstspan-
gselektronenmikroskop – **2** Abbildung
4.2.3.1.) von Silicium-(111)-Netzebenen –
chematische Darstellung zur Entstehung
(2) – **4** Aufnahme einer Zink-Aluminium-
erung mit dem Emissionselektronen-
oskop



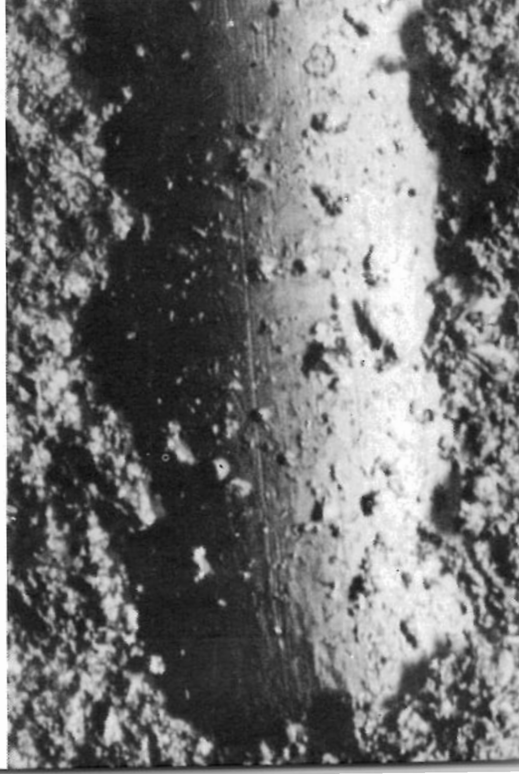


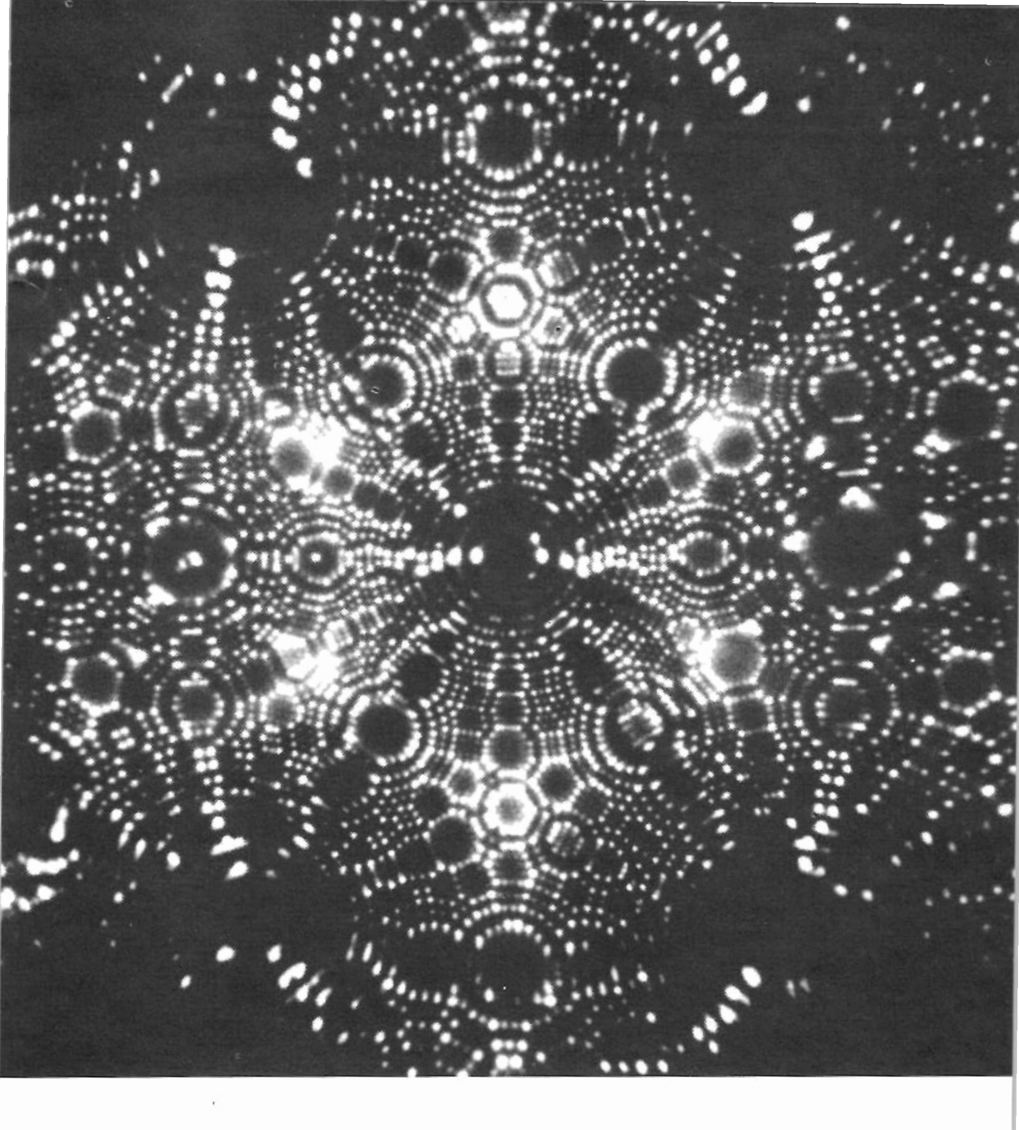
1 | 4
2 3 | 5 6

Tafeln 28 und 29 (↑ 1.1.3.5.) **1** Rasterelektronenmikroskop mit energiedispersivem Spektrometer – *1* elektronenoptische Säule, *2* Elektrik, *3* Kryostat mit Detektor und Vorverstärker, *4* Vielkanalanalysator mit Monitor (vgl. Abb. 1.1.3-10) – **2** bis **6** Rasterelektronenmikroskopische Aufnahmen eines aufgebrochenen Tantalelektrolytkondensators (Ta-Draht mit Sinterkörper); inverses Probenstrombild (**2**), Sekundärelektronenbild (**3**), Abbildung mit rückgestreuten Elektronen (RE) (**4**), **5** und **6** RE-Bild mit Festkörperdetektorenpaar (Verstärkung des Material- und Unterdrückung des Topographiekontrasts bzw. umgekehrt)



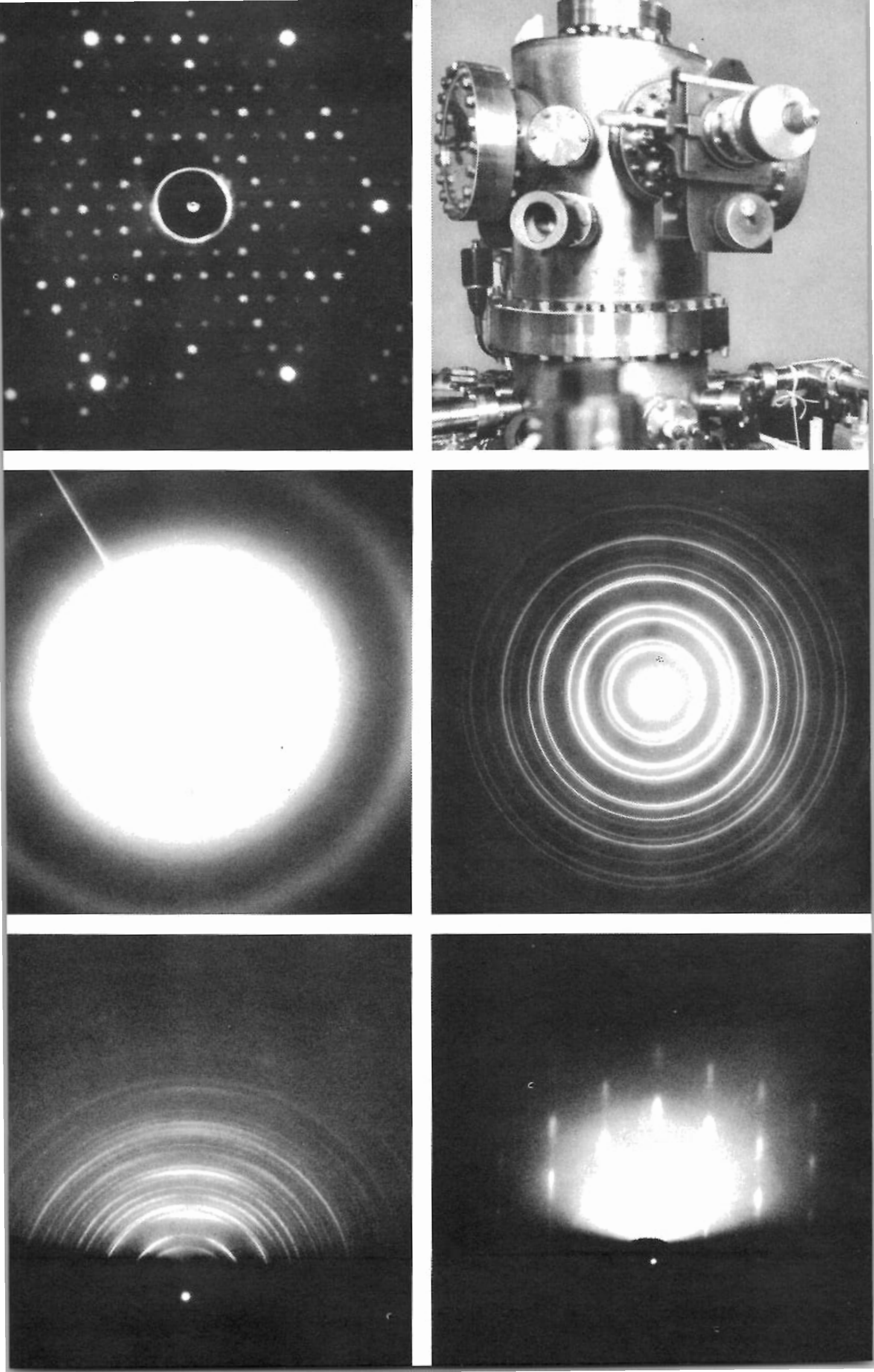


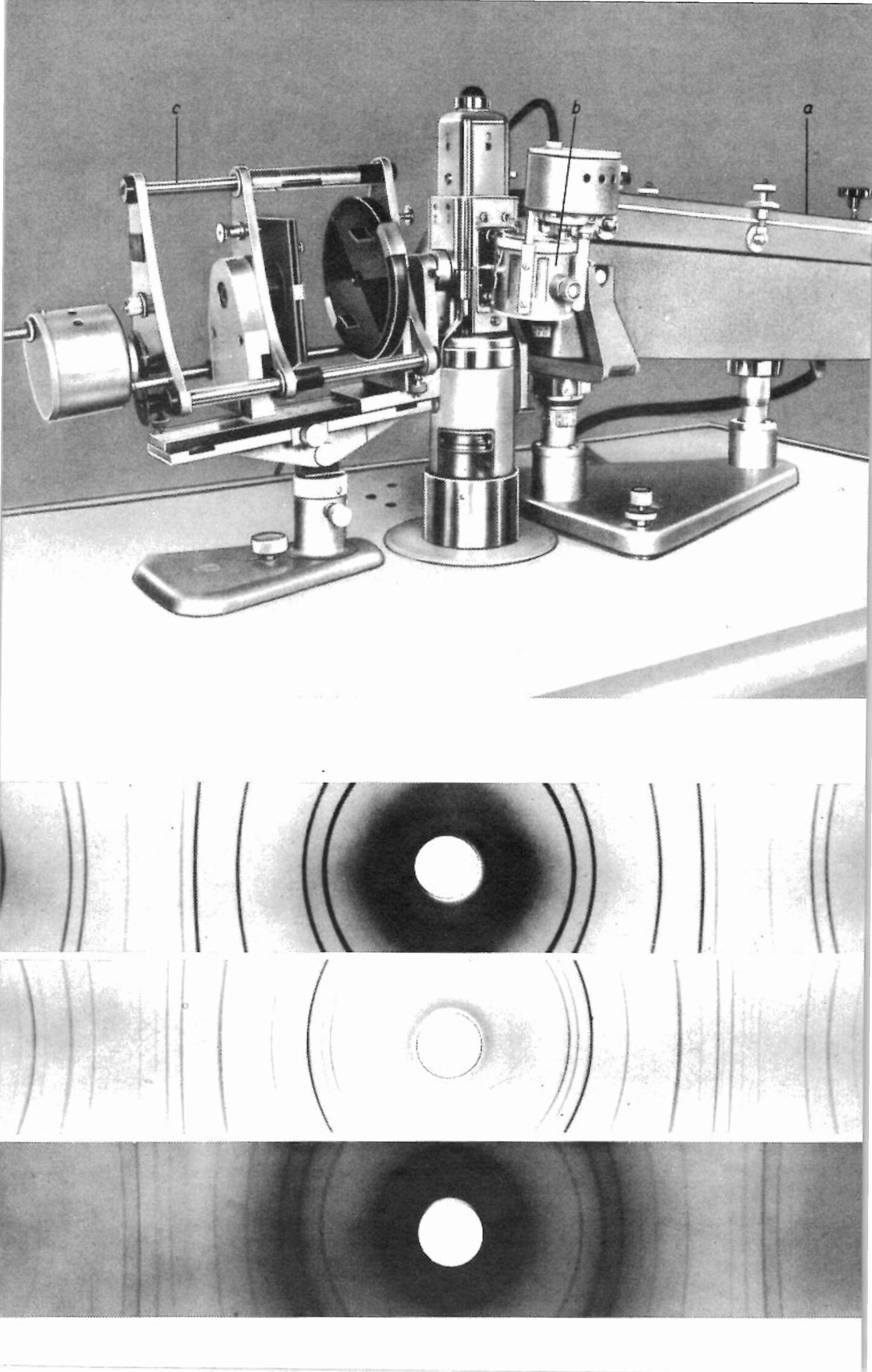




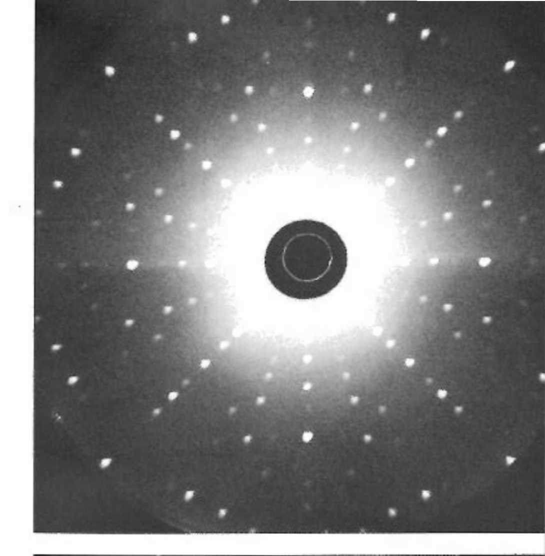
1	2 3
	4 5
	6 7

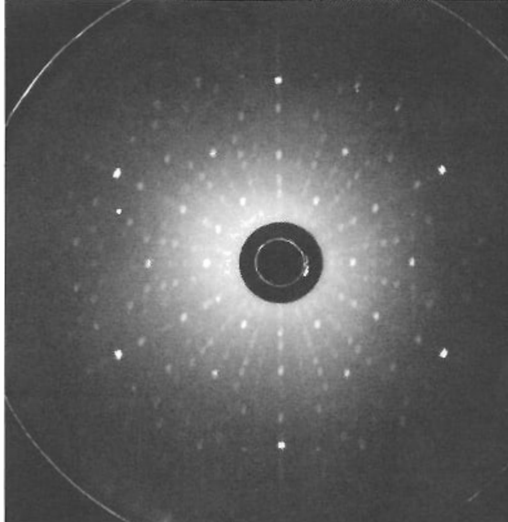
Tafeln 30 und 31 († 1.1.3.5., 1.2.3.3., 6.4.2.3.4.) **1** Feldionenmikroskopaufnahme einer Wolframkristallhalbkugel von 32 nm Radius – **2** Beugungsbild einer (111)-Siliciumoberfläche erhalten mit langsamen Elektronen (LEED); $U = 50$ V – **3** Ultrahochvakuumanlage zur Elektronenbeugung (LEED) – **4** Elektronenbeugungsbild einer amorphen PTFE-Schicht; $U = 35$ kV – **5** Elektronenbeugungsbild (Durchstrahlung) einer 26 nm dicken polykristallinen Goldschicht; $U = 50$ kV – **6** Reflexionsbeugungsbild einer Selenoberfläche (polykristallin mit Textur); $U = 50$ kV – **7** Reflexionsbeugungsbild einer einkristallinen GaAs-Schicht; $U = 50$ kV



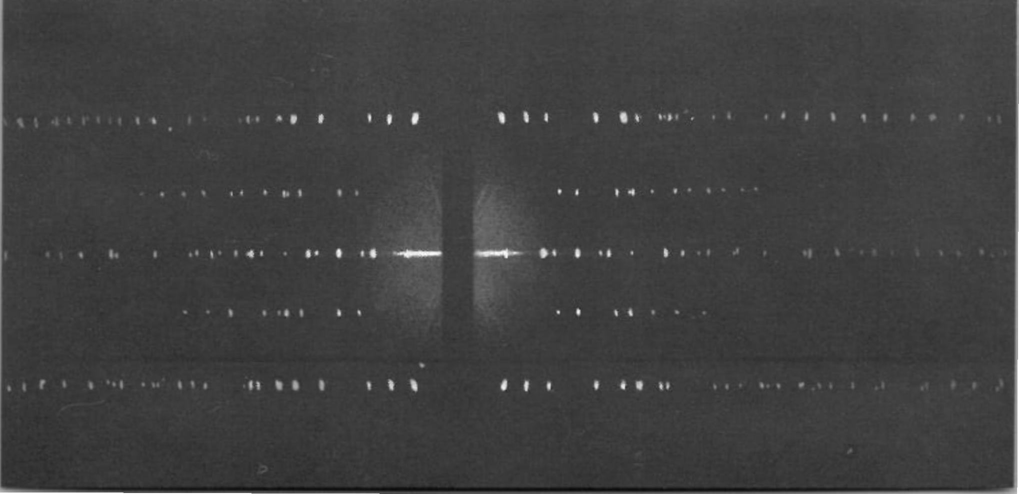

c
b
a

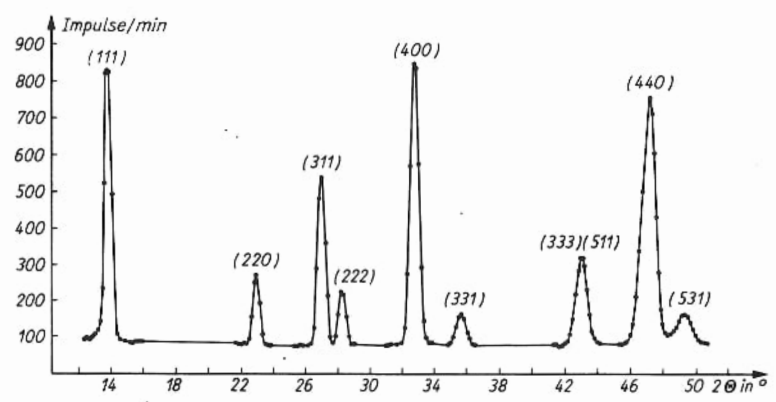
1 | 3
4
2 | 5





Tafeln 32 und 33 († 1.2.3.3., 6.4.2.3.) **1** Verschiedene an der Röhre eines Röntgengenerators angebrachte Kameras; *a* Röntgenkleinwinkelkamera, *b* Debye-Scherrer-Kamera, *c* Lauekamera – **2** Debye-Scherrer-Aufnahmen; *oben* von Aluminium (kubisch), *Mitte* von Zink (hexagonal), *unten* von Polytetrafluorethylen (partiell kristallin mit Kristallisationsgrad 65%) – **3** Lauebeugungsdiagramm von Spinell (kubisch) in [100]-Richtung – **4** Lauebeugungsdiagramm von Beryll (hexagonal) parallel zur *c*-Achse; aus den Diagrammen ist die vier- bzw. sechszählige Kristallsymmetrie der Proben erkennbar – **5** Drehkristallaufnahme von Marienglas (Gips, monoklin) parallel zur *c*-Achse

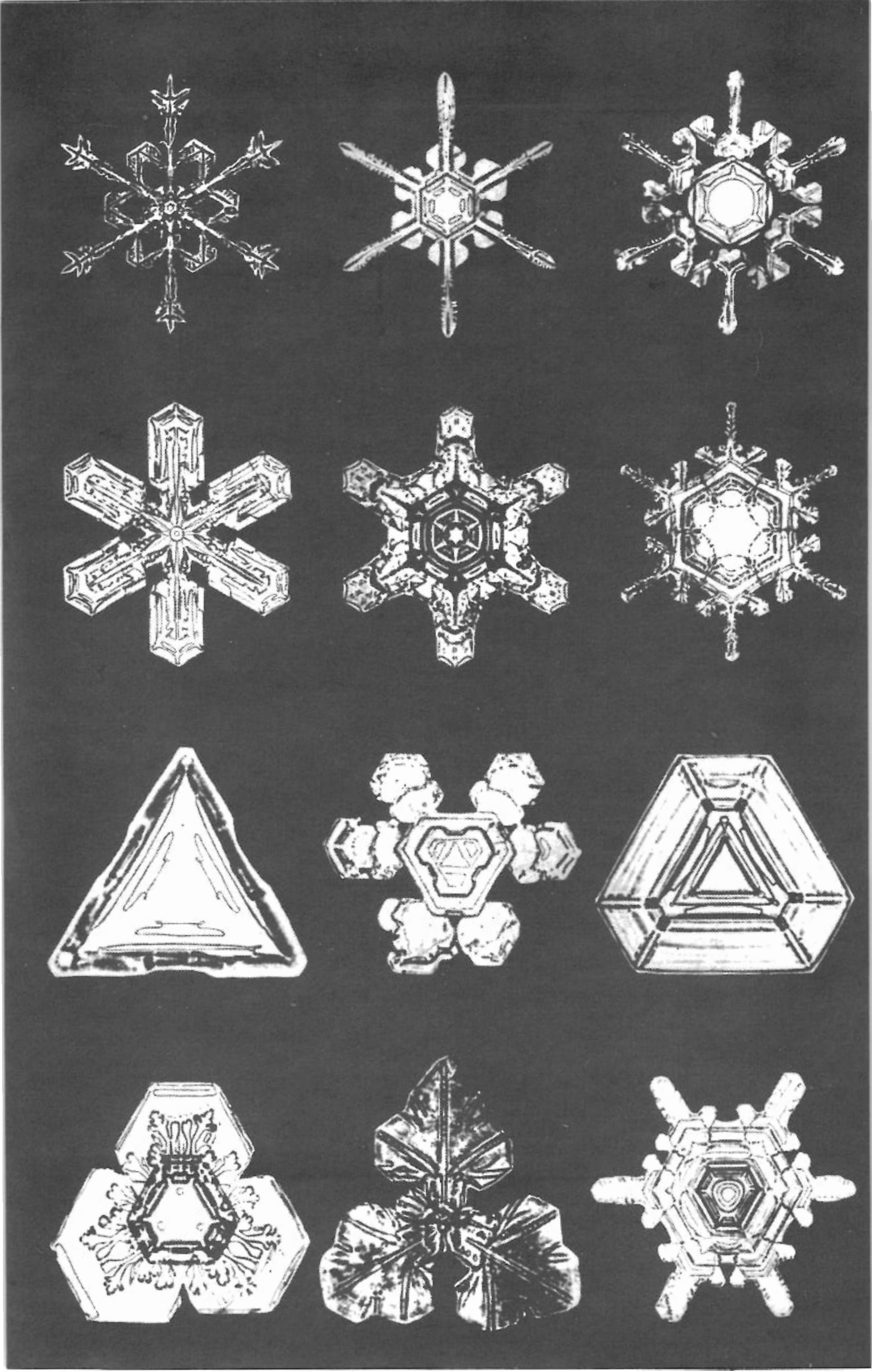




Impulse/min
900
800
700
600
500
400
300
200
100
(111)
(220)
(311)
(222)
(400)
(331)
(333)(511)
(440)
(531)
14
18
22
26
30
34
38
42
46
50
2Θ in °

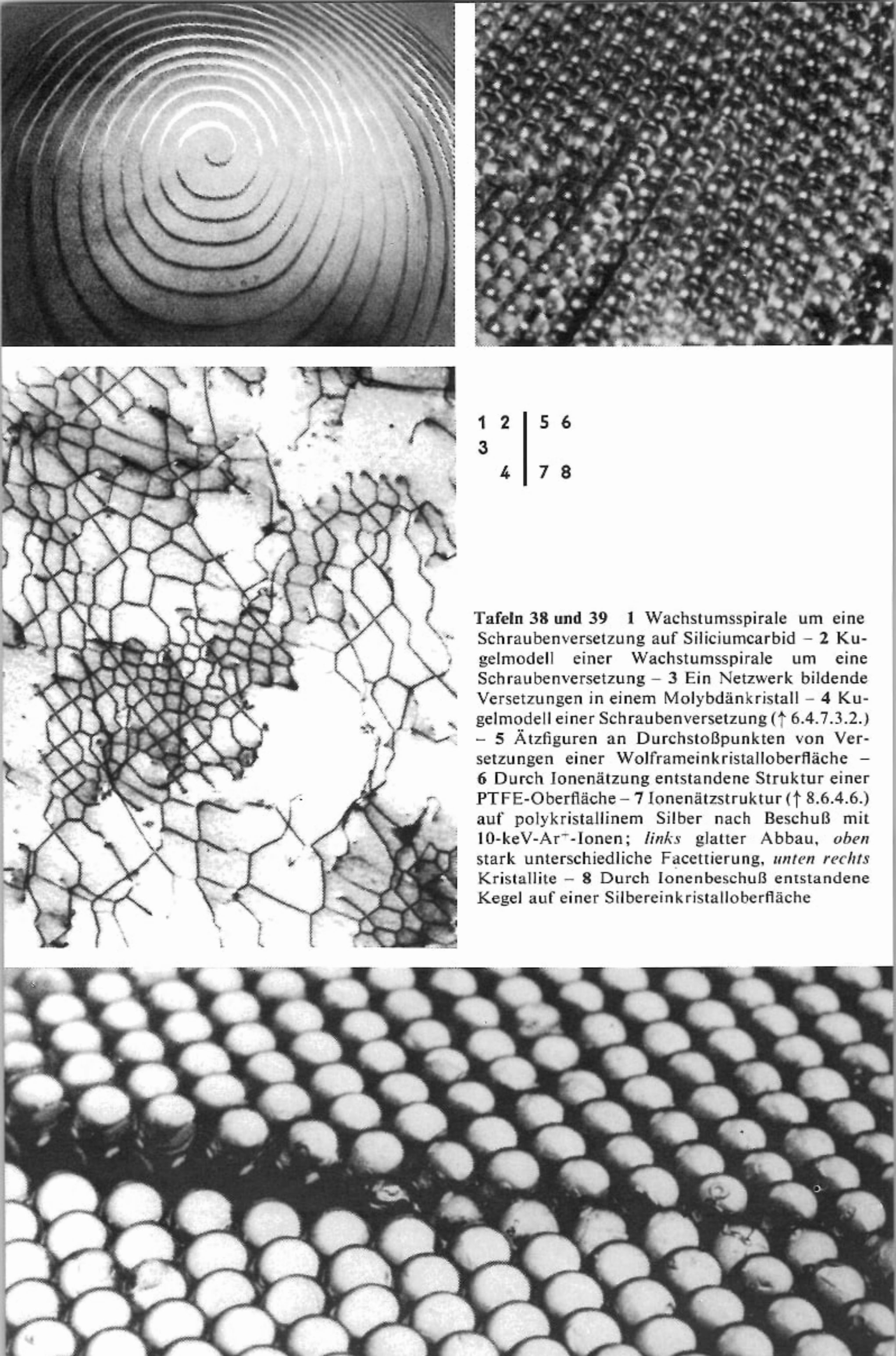
1	4
2	
3	5

Tafeln 34 und 35 (↑ 6.4.2.3.4.) **1** Neutronendiffraktometer am Reaktor des ZfK Rossendorf – **2** Neutronenbeugungsdiagramm (MnMg-Ferrit) – **3** Rechnergesteuertes Röntgenzählrohrdiffraktometer HZG 4 – **4** Elektronenoptische Aufnahme einer Glasoberfläche – **5** Kristallbildung in Glas



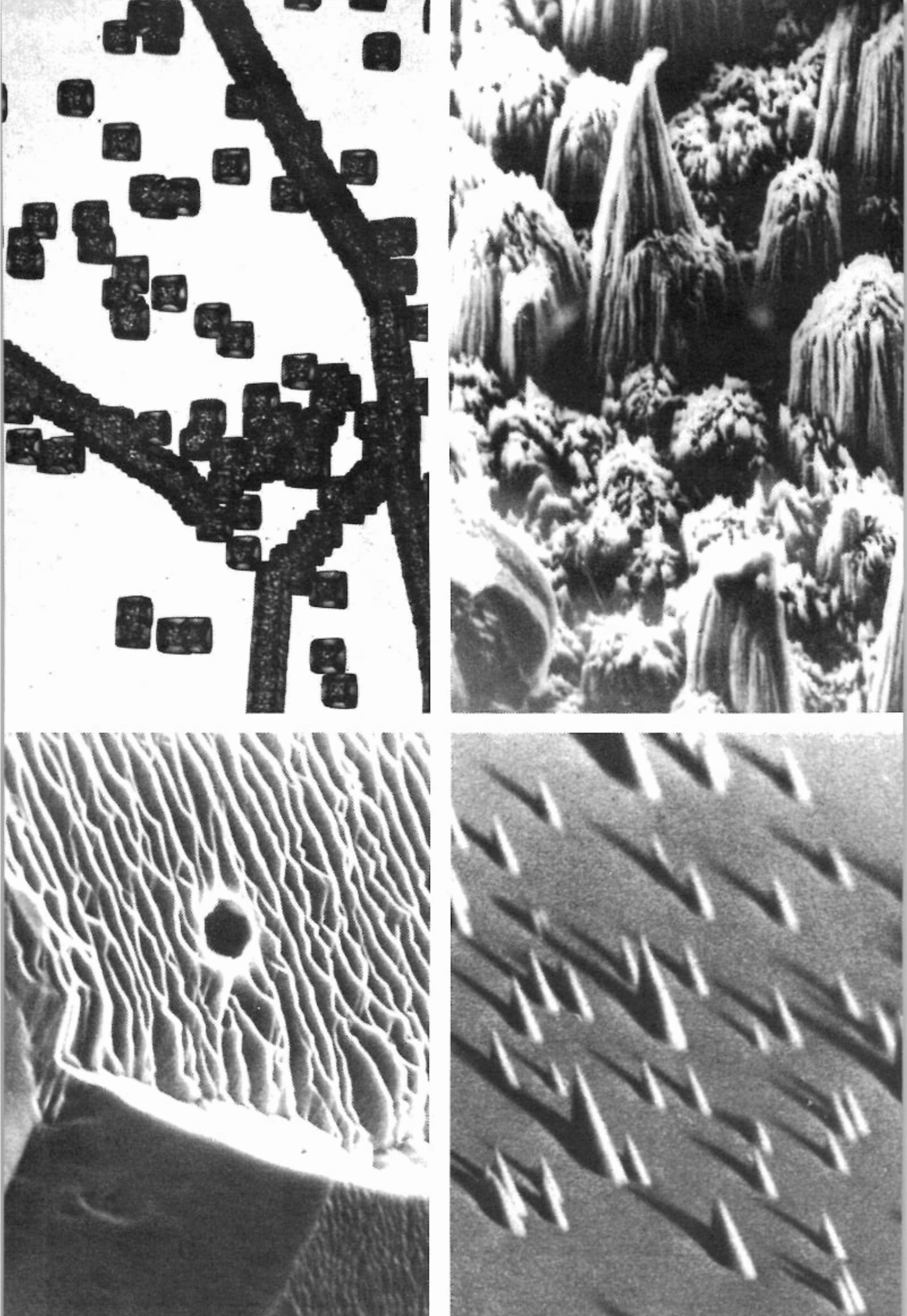
1 | 2 3
  | 4 5

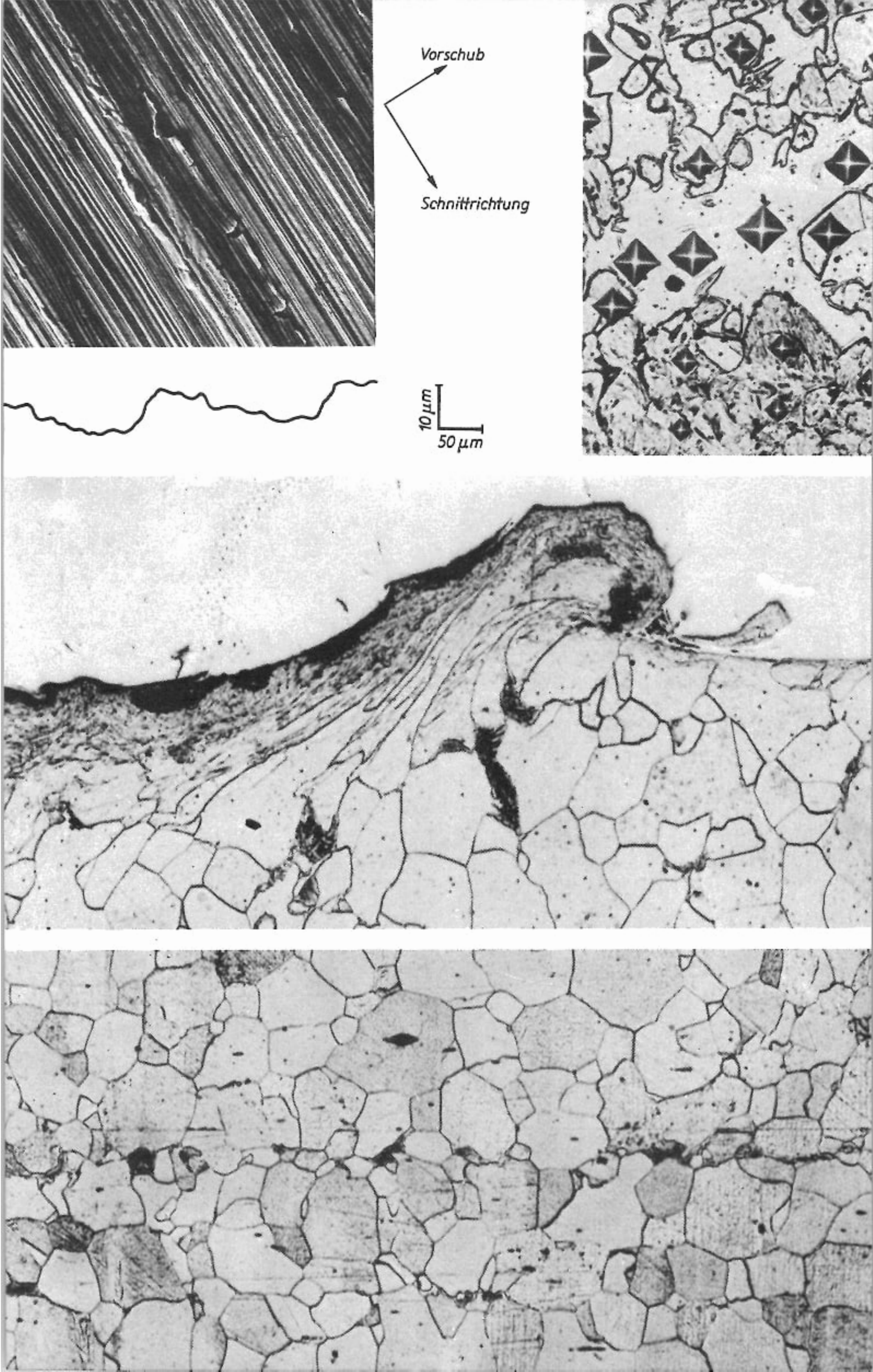
Tafeln 36 und 37 (↑ 6.4.) **1** Auswahl verschiedener Schneekristalle – **2** Oberfläche eines Platineinkristalls (kubisch, Vergrößerung 40000) – **3** Magnetit auf Eisen (kubisch, Vergrößerung 6000) – **4** ZnS-Schicht (hexagonal, Vergrößerung 6300) – **5** Cadmiumoxidschicht (kubisch, Vergrößerung 6300)

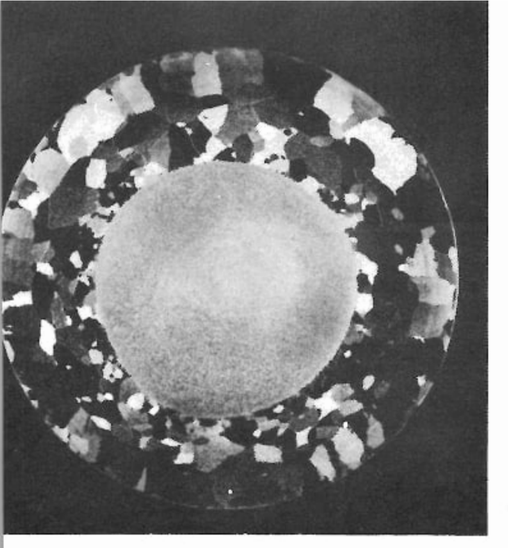


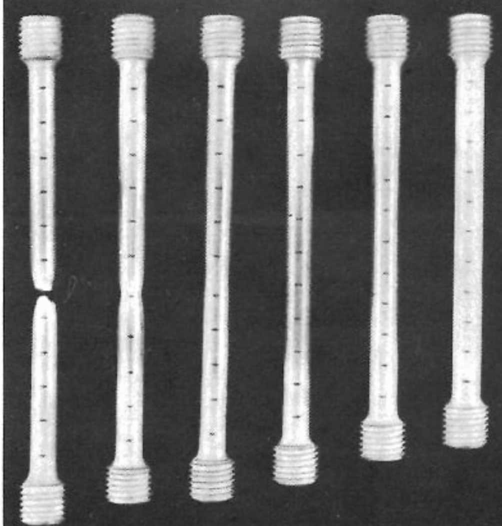
1 2 | 5 6
3
4 | 7 8

Tafeln 38 und 39 **1** Wachstumsspirale um eine Schraubenversetzung auf Siliciumcarbid – **2** Kugelmodell einer Wachstumsspirale um eine Schraubenversetzung – **3** Ein Netzwerk bildende Versetzungen in einem Molybdänkristall – **4** Kugelmodell einer Schraubenversetzung (↑ 6.4.7.3.2.) – **5** Ätzfiguren an Durchstoßpunkten von Versetzungen einer Wolframeinkristalloberfläche – **6** Durch Ionenätzung entstandene Struktur einer PTFE-Oberfläche – **7** Ionenätzstruktur (↑ 8.6.4.6.) auf polykristallinem Silber nach Beschuß mit 10-keV-Ar^+-Ionen; *links* glatter Abbau, *oben* stark unterschiedliche Facettierung, *unten rechts* Kristallite – **8** Durch Ionenbeschuß entstandene Kegel auf einer Silbereinkristalloberfläche



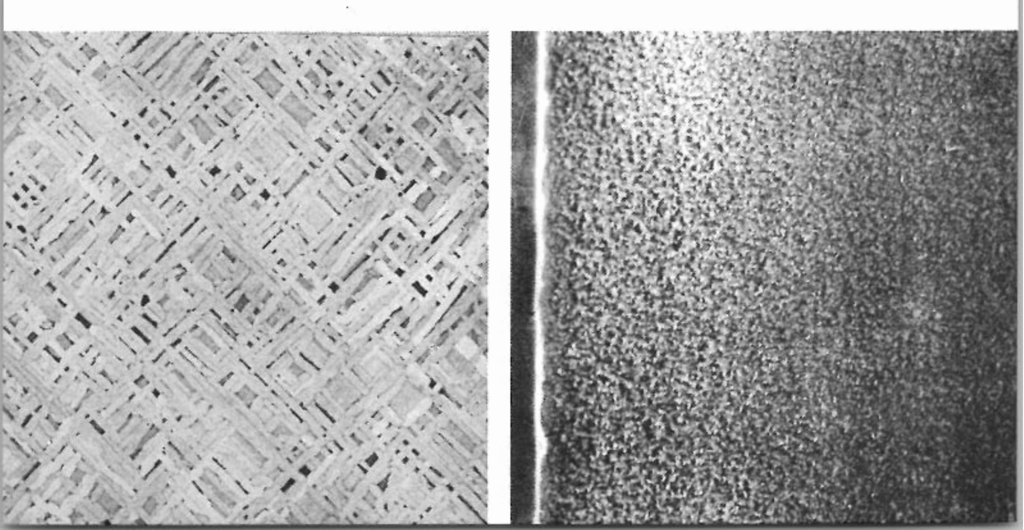

Vorschub
Schnittrichtung
10 µm
50 µm

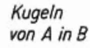




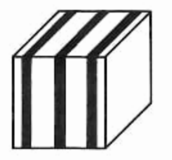
1 2	5 6
3	
4	7 8

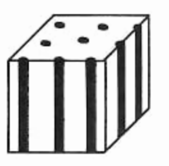
Tafeln 40 und 41 (↑ 6.4.8.1.2.) **1** Plastische Verformung an einer abgedrehten Stahlprobe (C 60) darunter Profil der dabei entstandenen Oberflächenrauhigkeit – **2** Stahl mit 0,3 % C von 740 °C in Wasser abgeschreckt (Ferrit und Martensit) mit Mikrohärteeindrücken – **3** Plastisches Fließen der Körner an einer durch Verschleiß aufgeworfenen Stahloberfläche – **4** Gefüge von Tiefziehblech – **5** Warmgezogene Aluminiumstange; durch kritische Verformung der Oberflächenschichten Grobkornbildung (Rekristallisation) – **6** Zerreißstäbe aus einer Aluminiumlegierung mit ursprünglich gleicher Länge nach unterschiedlichen Zugbelastungen – **7** Geätzter Anschliff eines in Grönland gefundenen Ni-reichen Eisenmeteoriten (Oktaedrit) mit typischer Entmischungsstruktur, den sogenannten Widmannstättenschen Figuren – **8** Querschliff einer durch Ionenstrahlnitrierung vergüteten Eisenoberfläche; *weißer Rand* reines Nitrit, *heller Bereich* Mischung Nitrit-Eisen (Dicke etwa 70 μm)

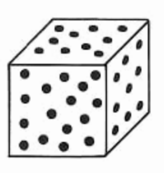




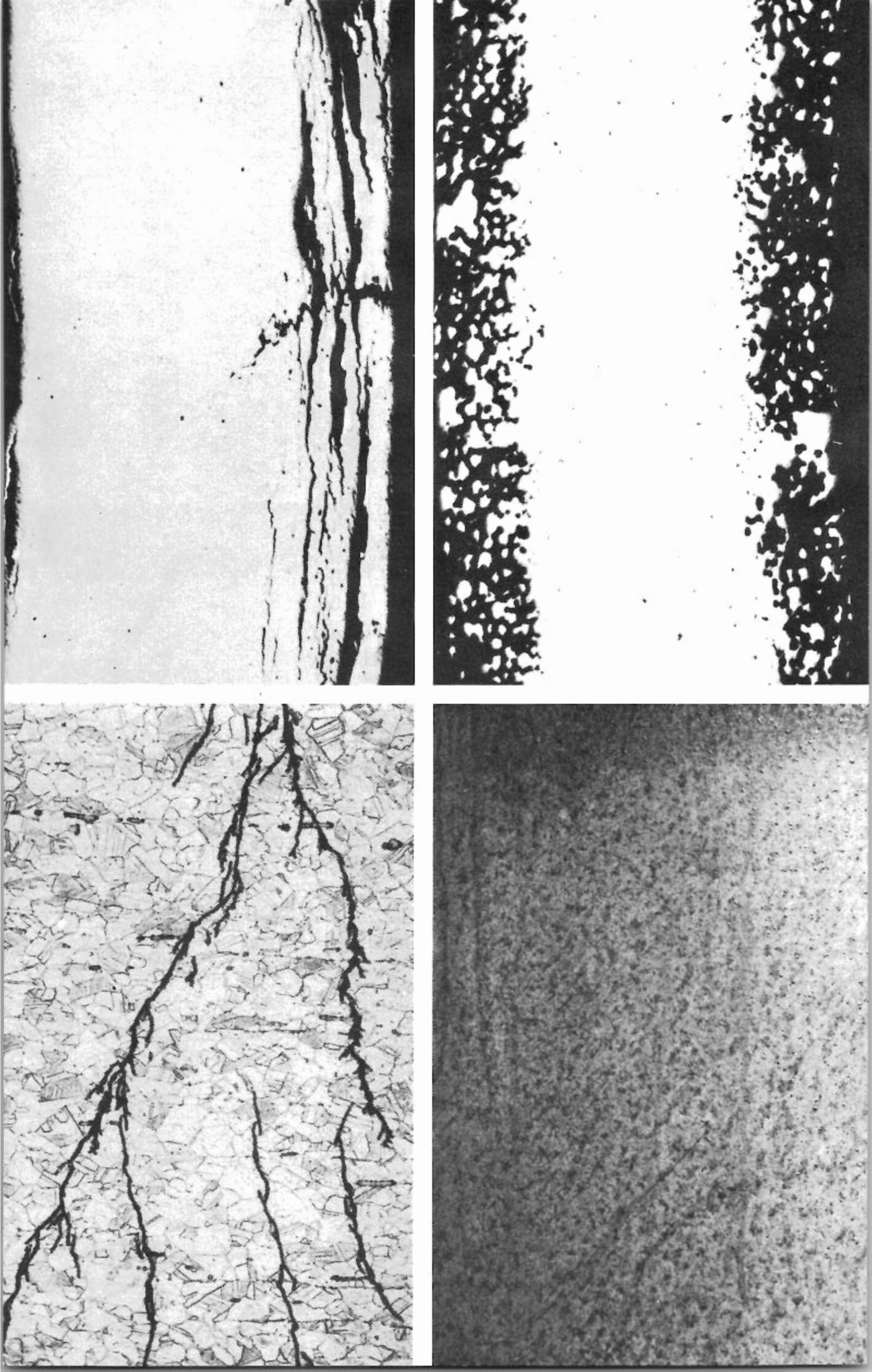






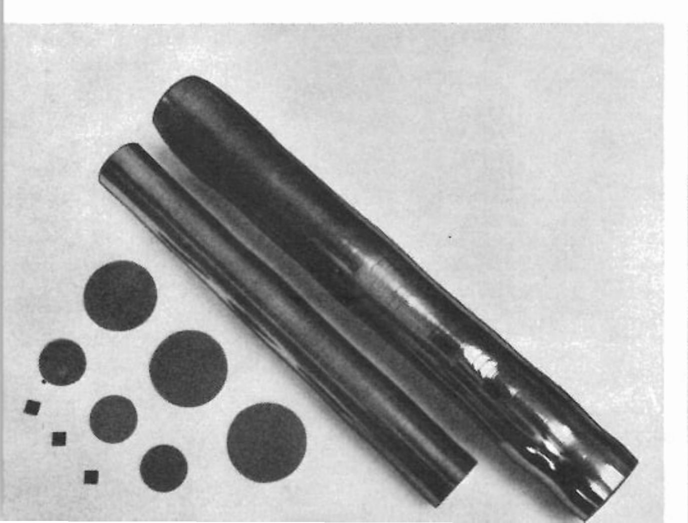
1	3 4
2	5 6

Tafeln 42 und 43 (↑ 6.5.7.) **1** Schematische Darstellung der Phasenstrukturen in Polymermischungen aus zwei Komponenten *A* und *B* – **2** Elektronenmikroskopische Aufnahmen von Filmen aus Styren-Isopren-Blockcopolymeren entsprechen weitgehend den in **1** dargestellten Strukturen; *links oben* 20% Styrenanteil (Polystyrenkugeln in einer Polyisoprenmatrix); *rechts oben* 40% Styrenanteil (Lamellenstruktur), *unten links* 60% Styren (stäbchenförmige bzw. zylindrische Verteilung des Polyisoprens in der Polystyrenmatrix); *unten rechts* 70% Polyisopren als Kugeln in der Polystyrenmatrix – **3** Schichtkorrosion und interkristalline Spannungsrißkorrosion bei einer AlCuMg-Legierung – **4** Interkristalline Korrosion bei einer AlMg-Legierung – **5** Transkristalline Spannungsrißkorrosion bei einem austenitischen Stahl – **6** Lochfraß an der Oberfläche einer Chemieausrüstung aus einer CrNiTi-Legierung



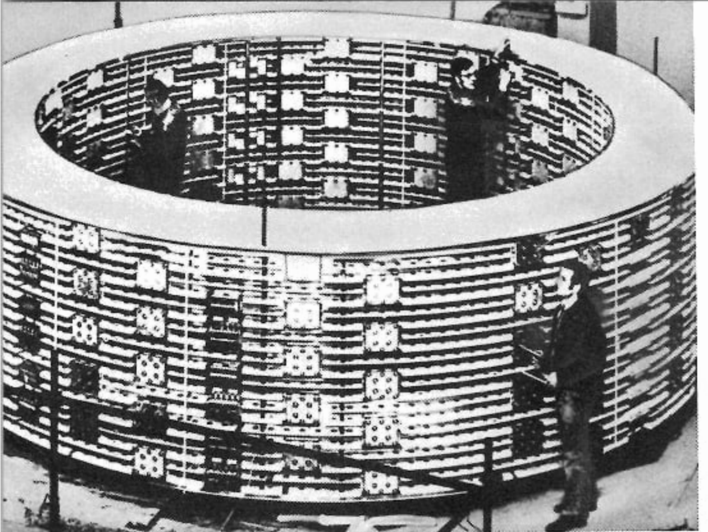
1 | 4 5
2 3 | 6

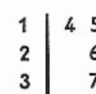
Tafeln 44 und 45 **1** Schliffbild eines legierten Leistungsgleichrichters; der als *schwarze Linie* im hellen Kristall sichtbare *pn*-Übergang ist vollkommen eben und entspricht somit den Anforderungen († 6.4.7.2.4., 8.3.) – **2** Siliciumeinkristalle in verschiedenen Bearbeitungsstadien – **3** Siliciumgleichrichter GY 125 (Gehäusekappe teilweise entfernt) – **4** Mikrocomputer auf einer einzigen Platine – **5** Elektronische Bauelemente (Dioden, Transistoren) und integrierte Schaltkreise – **6** Integrierter Schaltkreis mit mittlerem Integrationsgrad

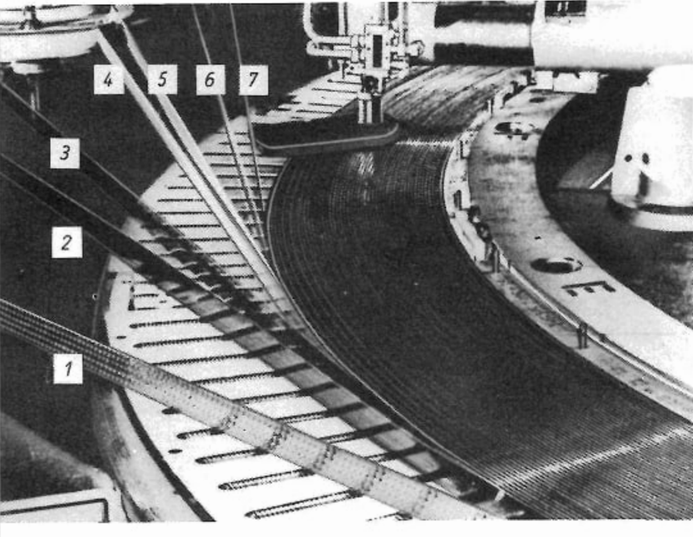




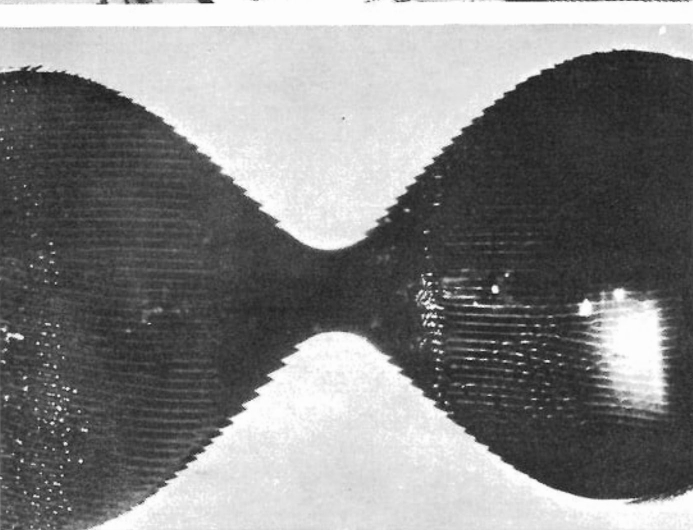
U253
IN
03A220

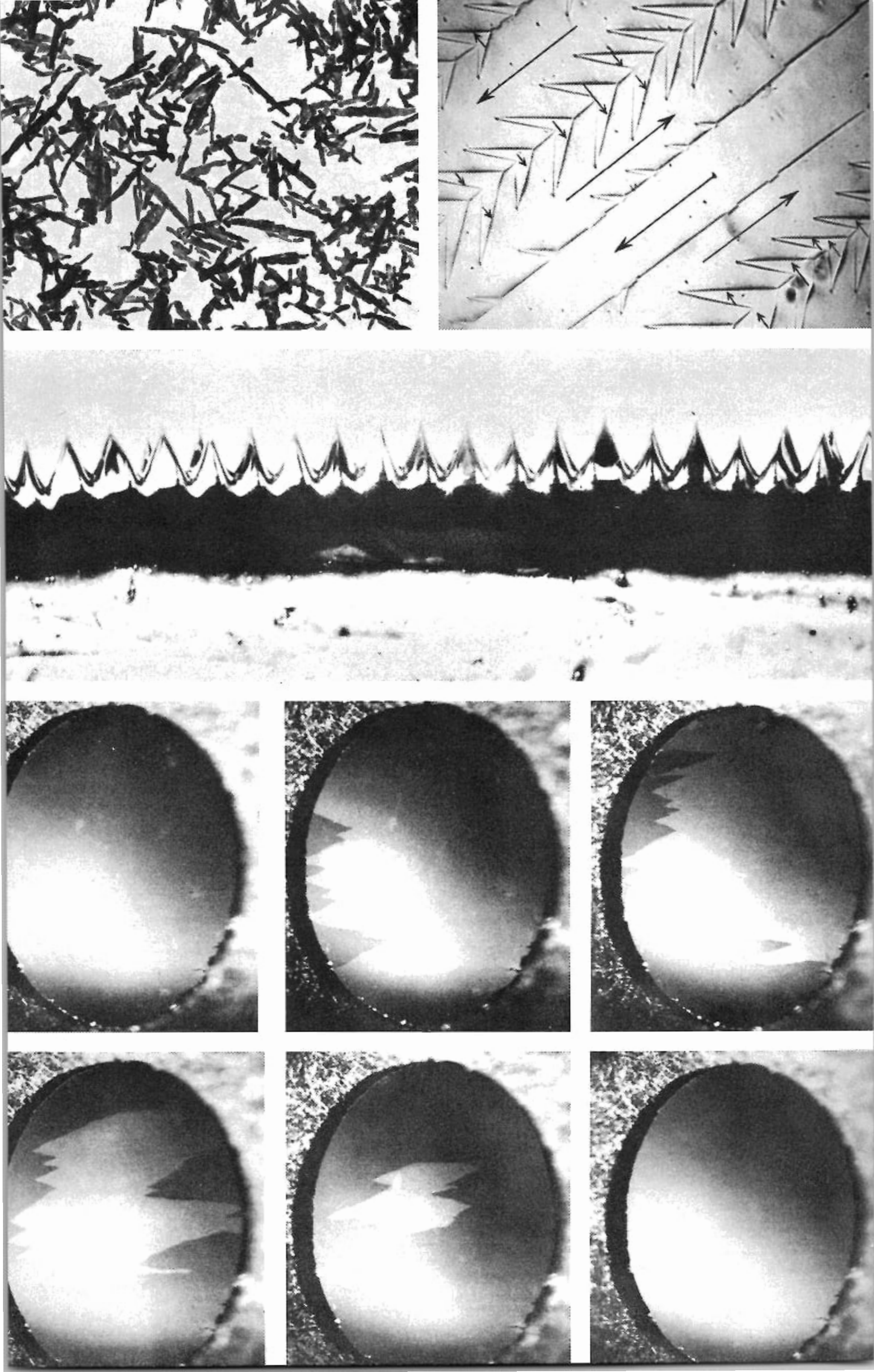




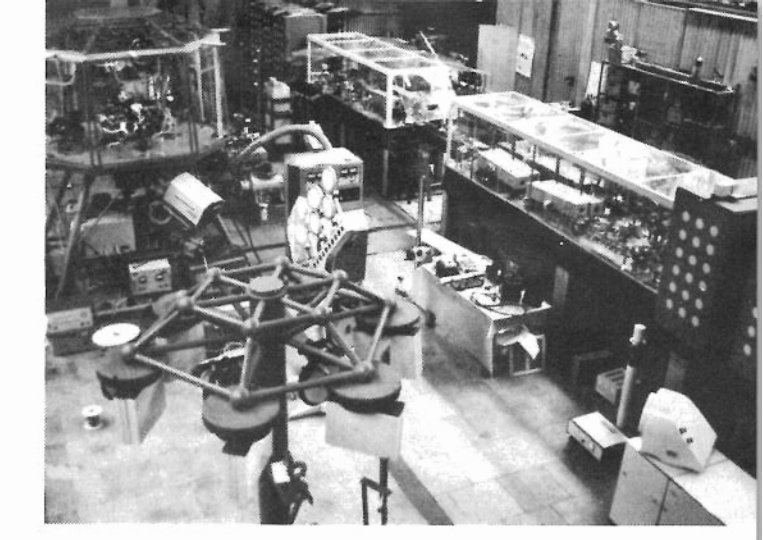


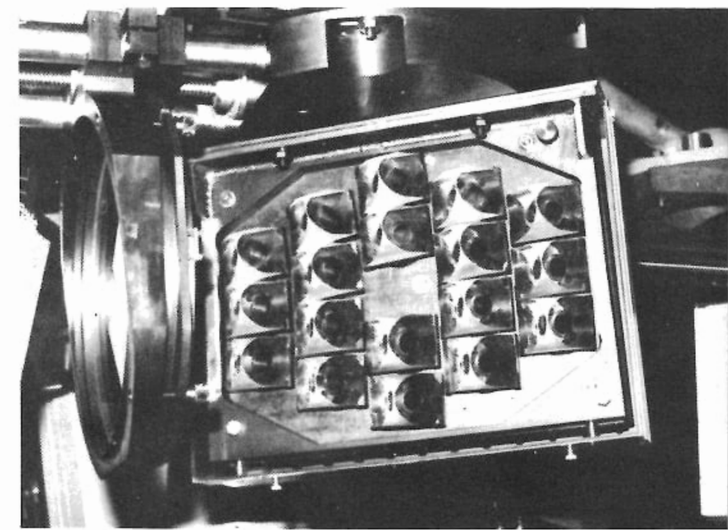
Tafeln 46 und 47 (↑ 8.5.4.1.)
1 Supraleitende Spule eines Blasenkammermagneten (BEBC) während der Montage; jede Spule besteht aus 20 einzelnen Spulen mit je 87 Windungen – **2** Wickeln einer Scheibenspule für den BEBC-Magneten; jede Scheibenspule besteht aus mehreren Bändern; *1* Kühlstreifen mit Öffnungen für den Durchfluß des flüssigen Heliums, *2* Stahlstreifen zur Stabilisierung, *3, 4, 6* Isolierstreifen, *5* Aluminiumband zur etwaigen Spulenerwärmung, *7* Supraleiter (NbTi-Legierung) in Kupfereinbettung – **3** Magnetische Flüssigkeiten hoher Stoffkonzentration; die Anziehung der Magnetpole ist so stark, daß die Flüssigkeit frei im Raum zwischen den Polen schwebt – **4** Elektronenmikroskopische Aufnahme der CrO_2-Schicht eines Videomagnetbands (10000fache Vergrößerung) – **5** Weißsche Bezirke (↑ 8.7.3.5.) in einem Einkristall aus siliciumhaltigem Eisen (die *Pfeile* geben die Magnetisierungsrichtung an) – **6** Wirkung eines Magnetfelds senkrecht zur Oberfläche einer magnetischen Flüssigkeit; die flüssigen Spitzen machen den Verlauf der Magnetfeldlinien sichtbar – **7** Stufenweise Ummagnetisierung einer Mo-Permalloy-Schicht durch ein schrittweise verstärktes Magnetfeld

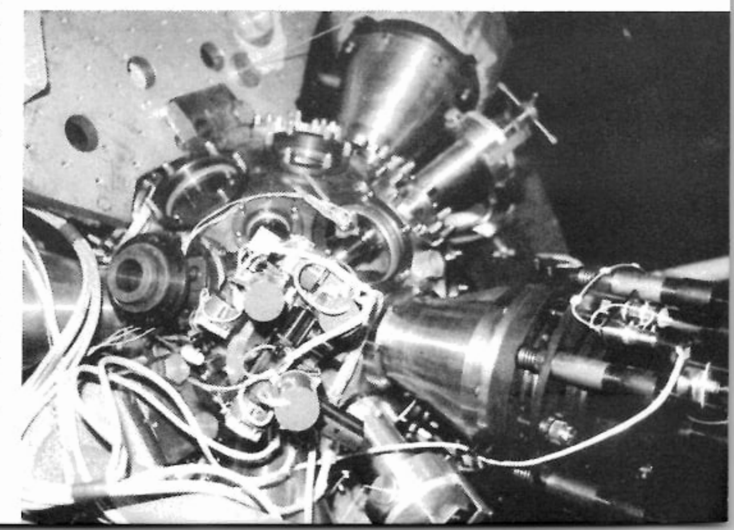




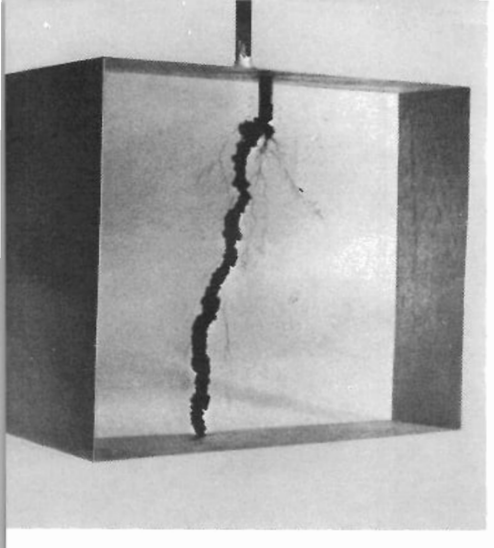
1	3
	4
2	5

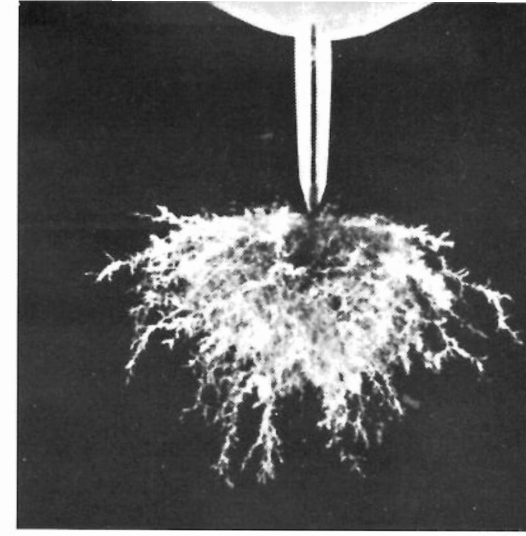


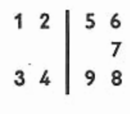




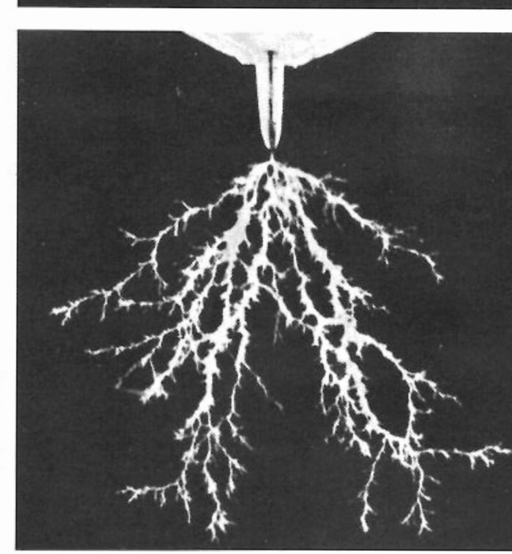
Tafeln 48 und 49 (↑ 9.3.)
1 600-t-Hochdruckpresse, mit der durch geeignete Hochdruckkammern so extrem hohe Drücke erzeugt werden können, wie sie z. B. im Erdinnern herrschen – **2** Bedampfungsanlage B 63 zur Materialbeschichtung – **3** Sowjetische Kernfusionsversuchsanlage „Delphin"; *oben links* Reaktionskammer, *unten rechts* einer der Endverstärker mit je 18 Lasergeräten, *vorn links* Multiprismenspiegel – **4** Multiprismenspiegel zur Umlenkung von je 18 parallelen Laserlichtbündeln; *links* Vorderlinse – **5** Reaktionskammer für das Fusionstarget mit Öffnungen für den Laserstrahleintritt

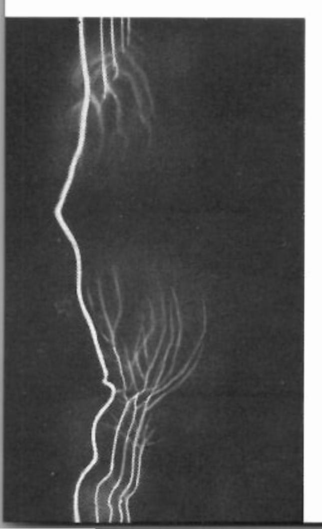


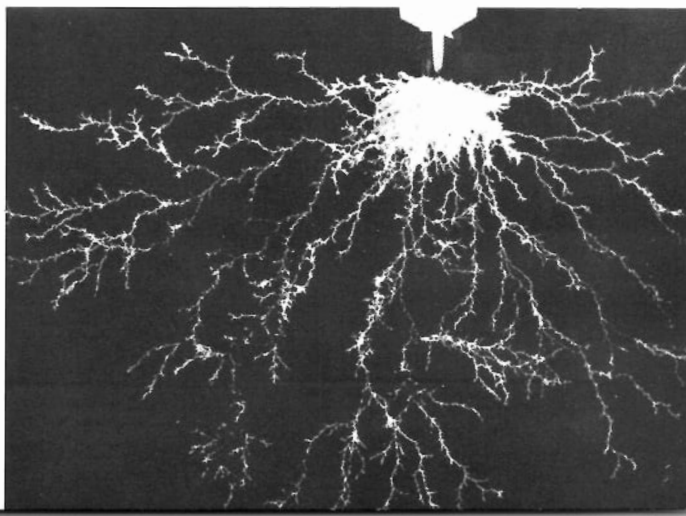




Tafeln 50 und 51 (↑ 9.1.6.2.) **1** Blitze über dem Hafen von Helsinki – **2** Kugelblitz – **3, 4** Funkenentladungen bei Spannungen von 0,1 und 2,25 MV – **5** Piacrylprüfkörper nach dem Durchschlag mit feinen verästelten Teilentladungskanälen in der Umgebung des Durchschlagkanals – **6, 7, 8** Teilentladungskanäle an einer Nadelspitze in Piacryl – **9** Funkenentladung mit schnell bewegter Platte (nach WESTPHAL)

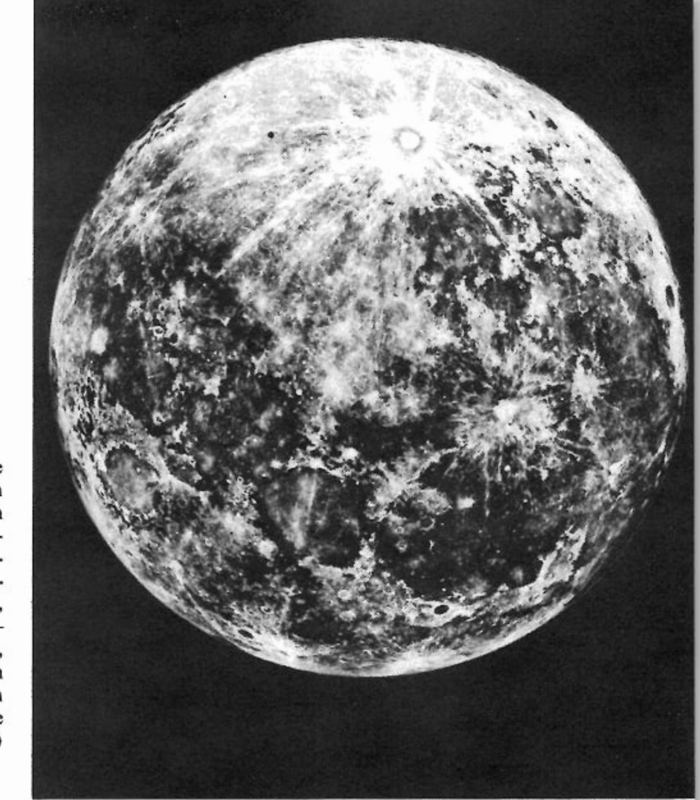


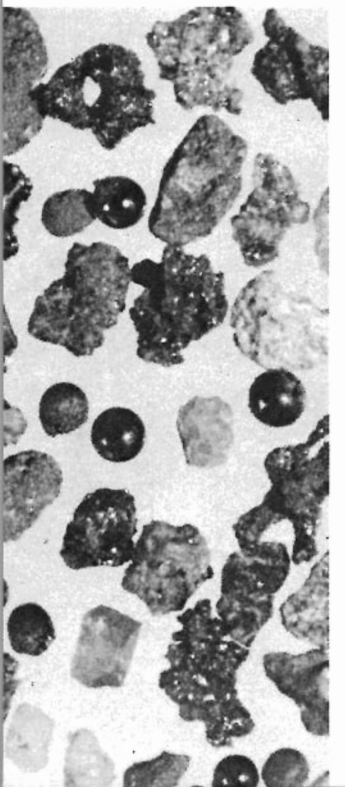


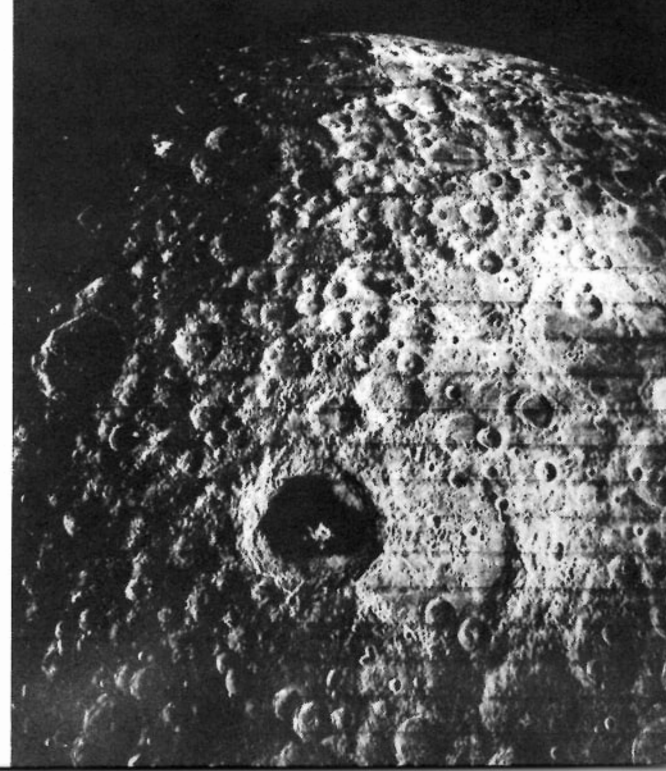


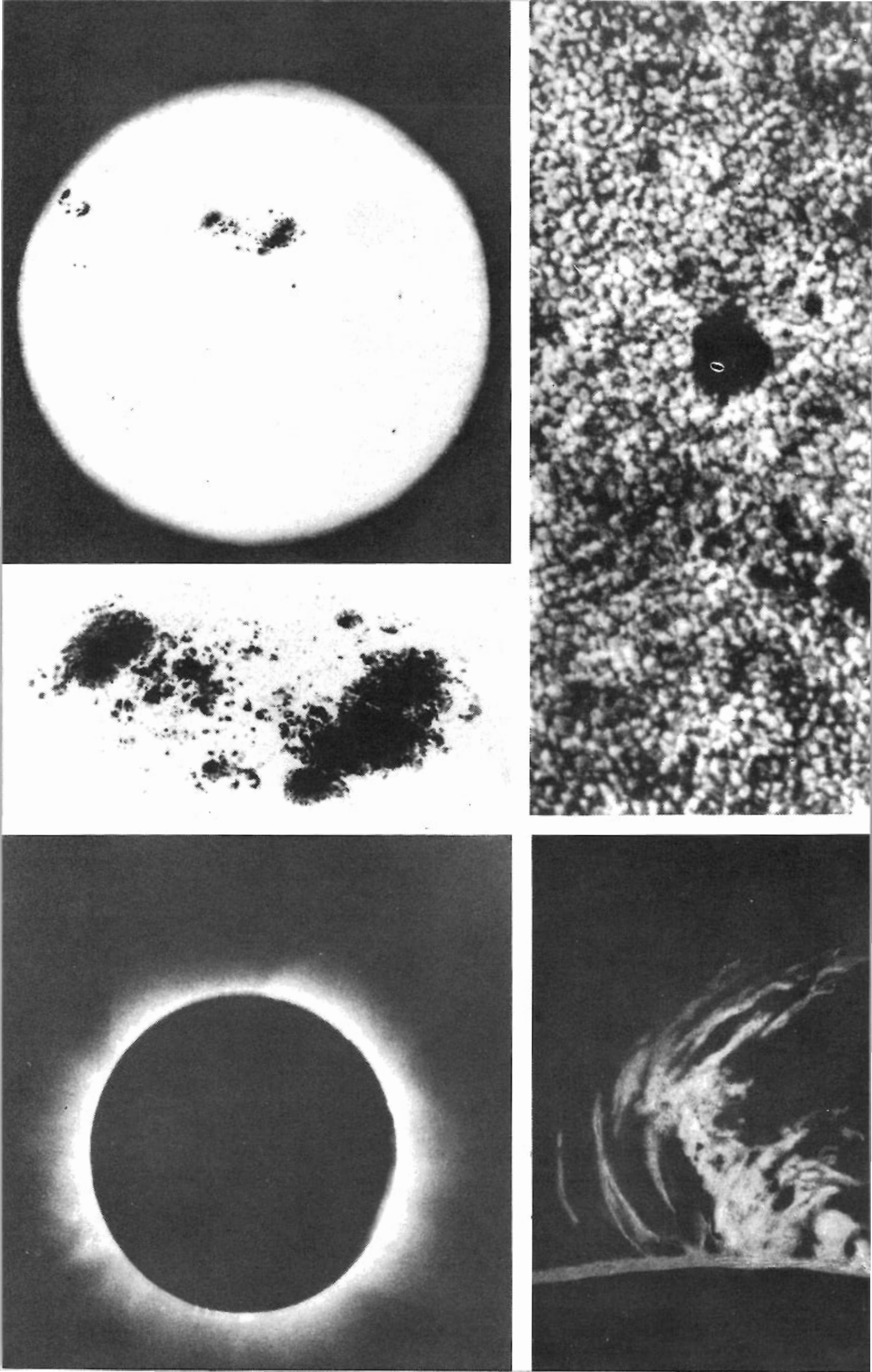
1 | 3
2 | 4 5

Tafeln 52 und 53 **1** Aufnahme der Erde aus dem Weltraum (Gemini-10-Kapsel über Indien und Sri Lanka, 1966) – **2** Infrarotaufnahme (↑ 10.1.2.2.2.) mittels eines Wettersatelliten (Mittelmeer, Italien, Griechenland, Balkanländer und Alpenkette) – **3** Mondoberfläche mit Kratern, Bergen, Rillen sowie Ebenen (↑ 10.2.1.) – **4** Mondgestein von Lunochod – **5** Mondrückseite (Aufnahme von Lunar Orbiter 3)



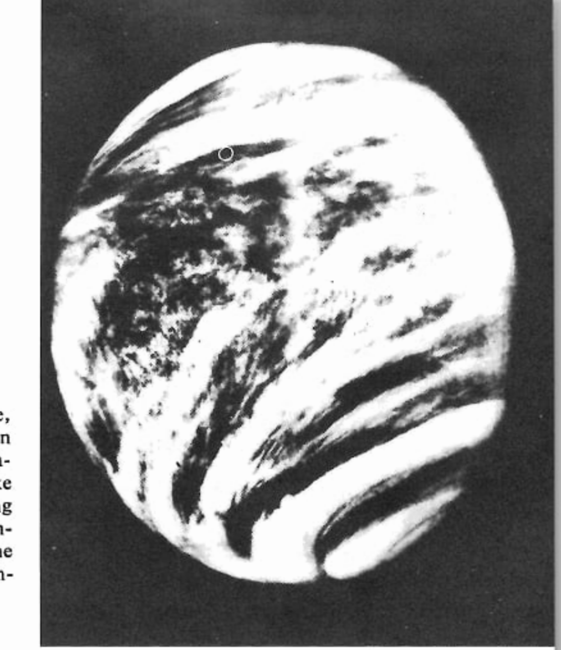


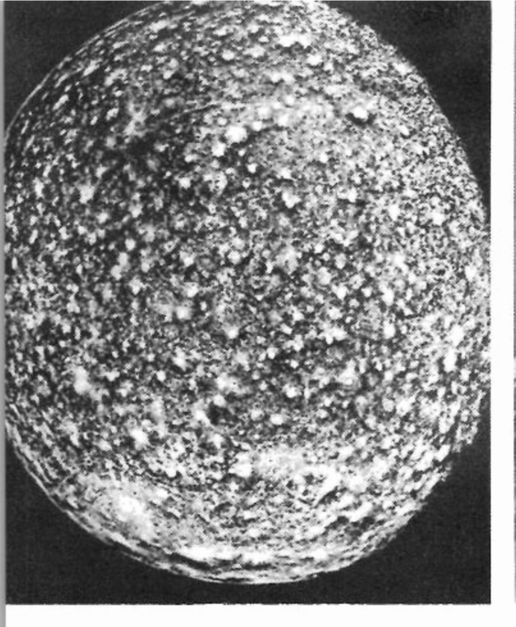




1 2 | 5
3 4 | 6
| 7 8

Tafeln 54 und 55 **1** Sonne mit Fleckengruppe, darunter vergrößert (↑ 10.2.2.) – **2** Granulation der Sonne – **3** Sonnenfinsternis – **4** Sonnenprotuberanz – **5** Venus (↑ 10.2.3.4.) mit Wolkendecke (Mariner 10) – **6** Panoramabild der Umgebung der Landestelle von Venus 10 (der *Pfeil* bezeichnet den Dichtemesser) – **7** Merkuroberfläche (↑ 10.2.3.3.) mit einem Krater von 12 km Durchmesser (Mariner 10) – **8** Merkur (Mariner 10)



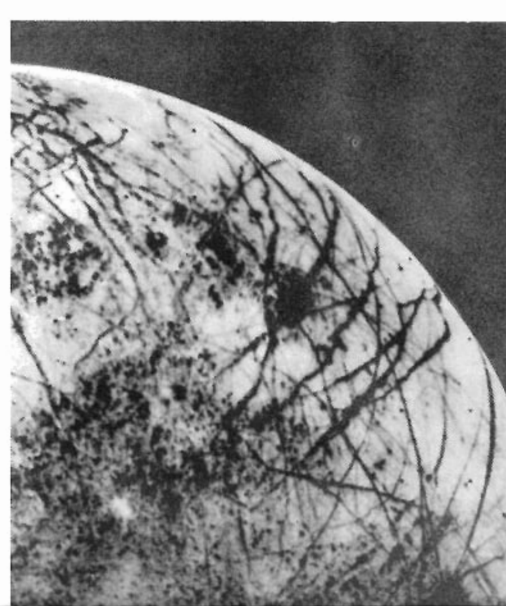


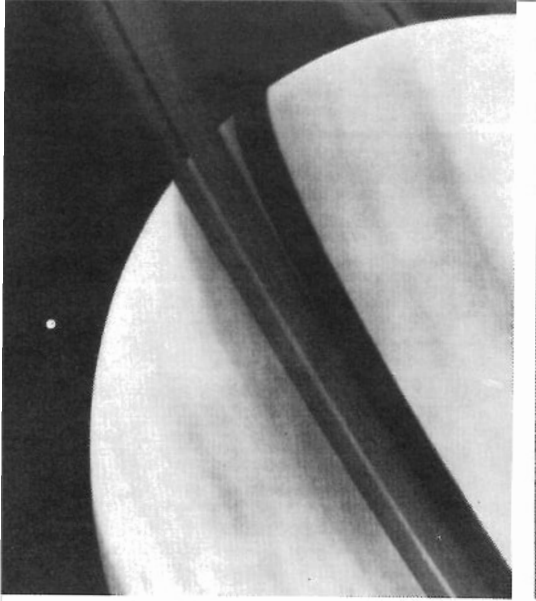


1 | 3 4
2 | 5 6

Tafeln 56 und 57 **1** Jupiter (↑ 10.2.3.6.) mit zwei Monden (Voyager II) – **2** Jupiter mit „rotem Fleck“ (Voyager I) – **3** Jupitermond Kallisto (Voyager II) – **4** Jupitermond Ganymed (Voyager I) – **5** Jupitermond Io (Voyager I) – **6** Jupitermond Europa (Voyager II)



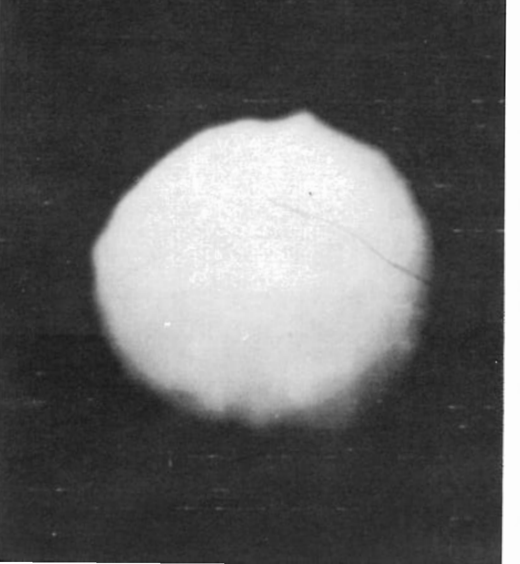


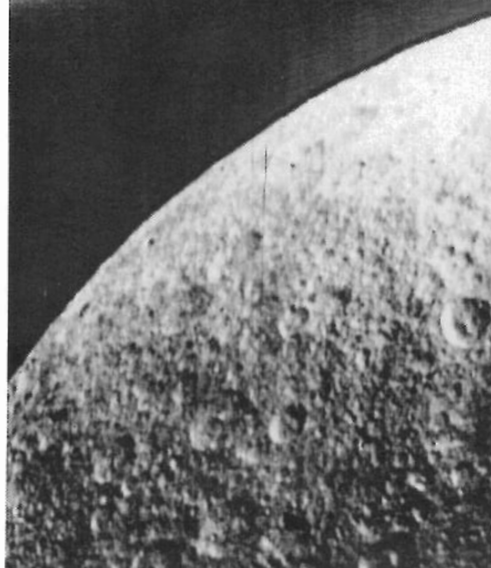


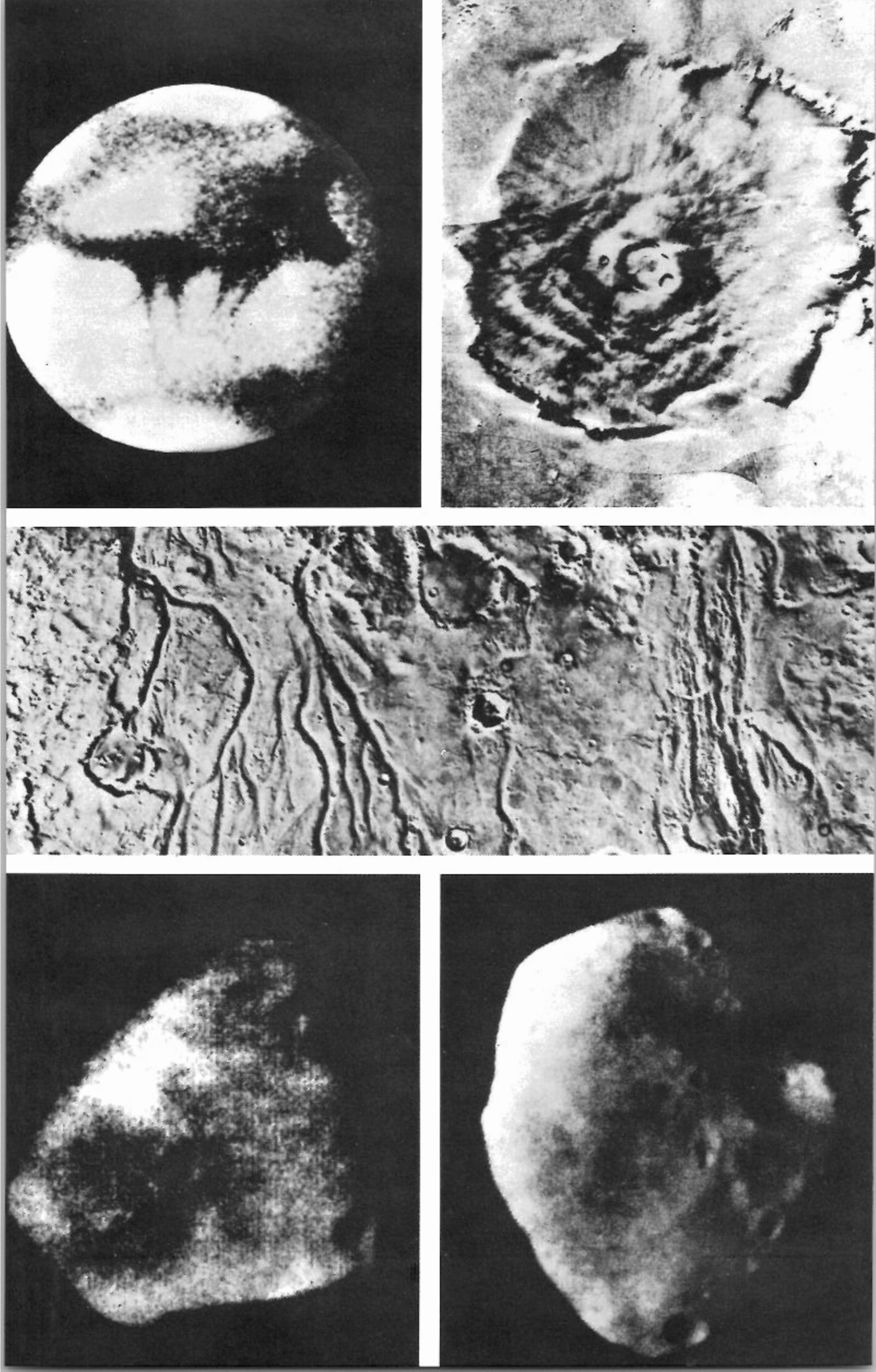


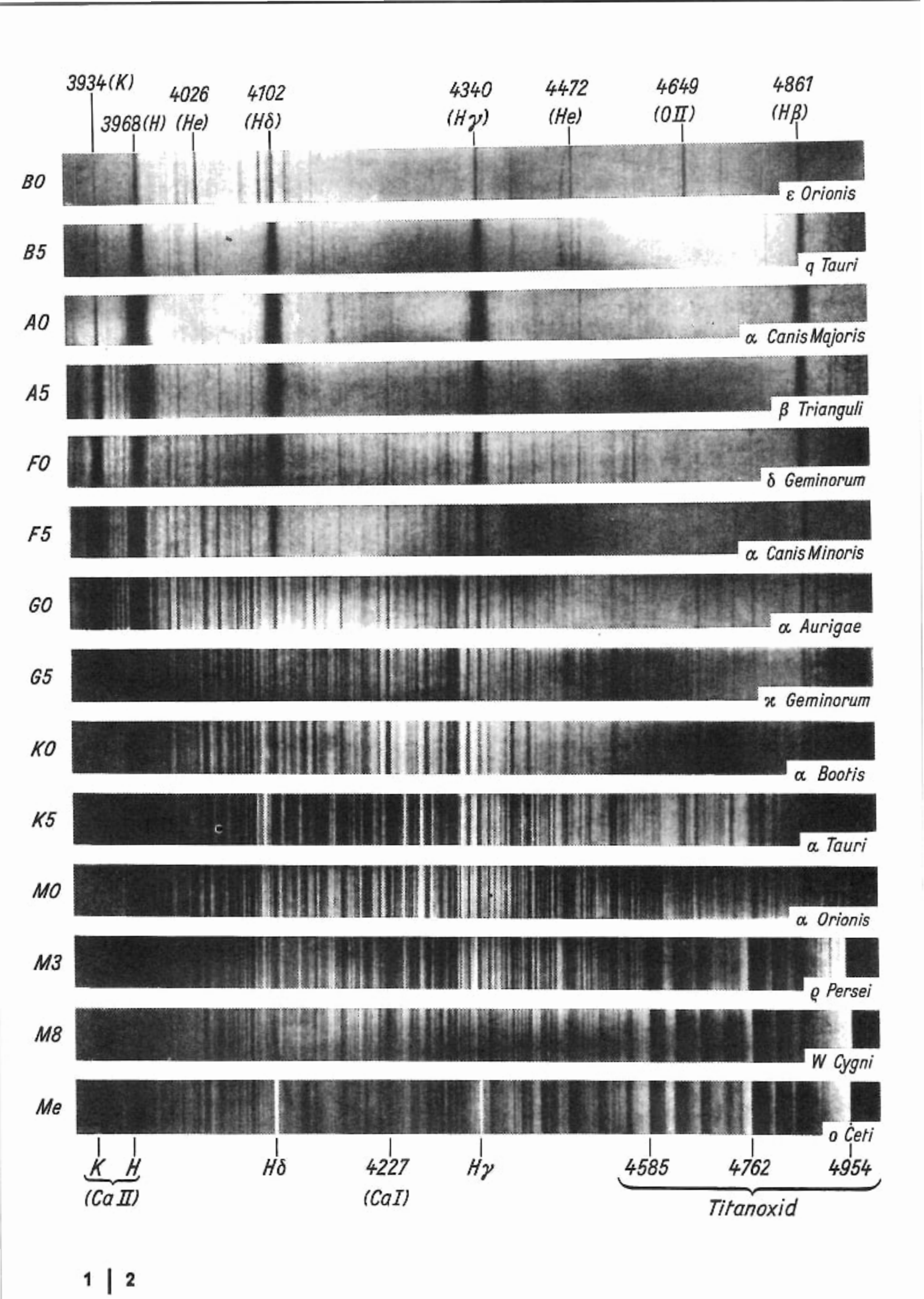
1 2 | 5 6
7
3 4 | 8 9

Tafeln 58 und 59 **1** Saturn mit Mond Tethys (↑ 10.2.3.7.); die Lücke zwischen den beiden Ringen beträgt etwa 3500 km (Voyager 1) – **2** Saturnoberfläche mit nordpolaren Gebieten; die hellen Details in der oberen Bildhälfte sind Wolkenbewegungen (Voyager I) – **3** Saturnmond Titan, der als einziger Trabant unseres Sonnensystems über eine eigene Atmosphäre verfügt (Pioneer 11) – **4** Saturnmond Rhea, die größeren Krater haben einen Durchmesser bis zu 75 km (Voyager I) – **5** Mars (↑ 10.2.3.5.) – **6** Marsvulkan Olympus Mons, der mit 600 km Durchmesser und etwa 23 km Höhe der wahrscheinlich größte Schildvulkan des Planetensystems ist (Mariner 9) – **7** Marsflußbetten – **8** Marsmond Deimos (Mariner 9) – **9** Marsmond Phobos (Mariner 9)



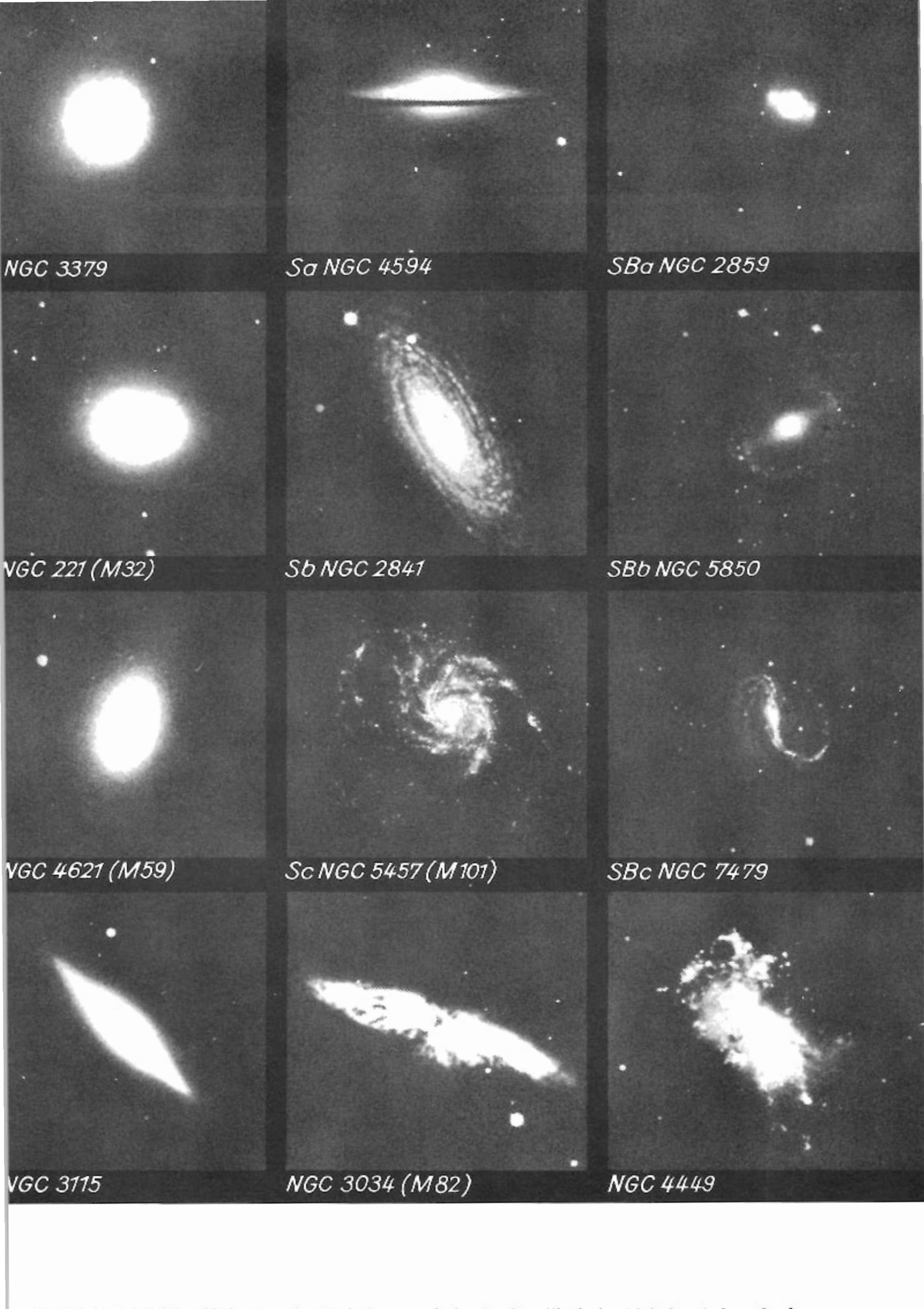






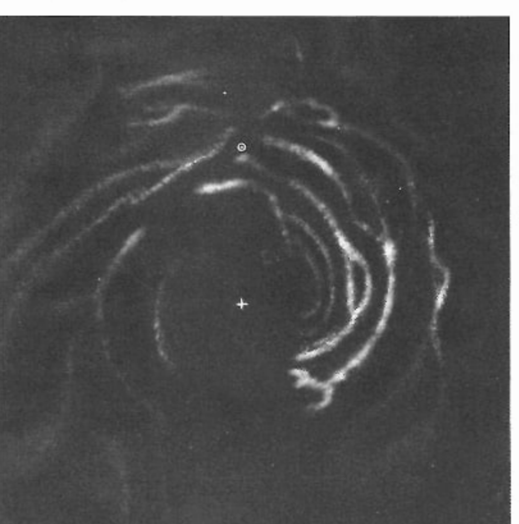
1 | 2

Tafeln 60 und 61 1 Sternspektren verschiedener Spektralklassen; für einige Linien sind die Wellen in 10^{-1} nm, die Namen und die sie erzeugenden Elemente angegeben (↑ 10.3.1.3.) –



2 Hubblesche Klassifizierung der Nebeltypen; *linke Spalte* elliptische Nebel, *mittlere Spalte* normale Spiralnebel und irregulärer Nebel NGC 3034, *rechte Spalte* Balkenspiralen und irregulärer Nebel NGC 4449 (↑ 10.4.1.)

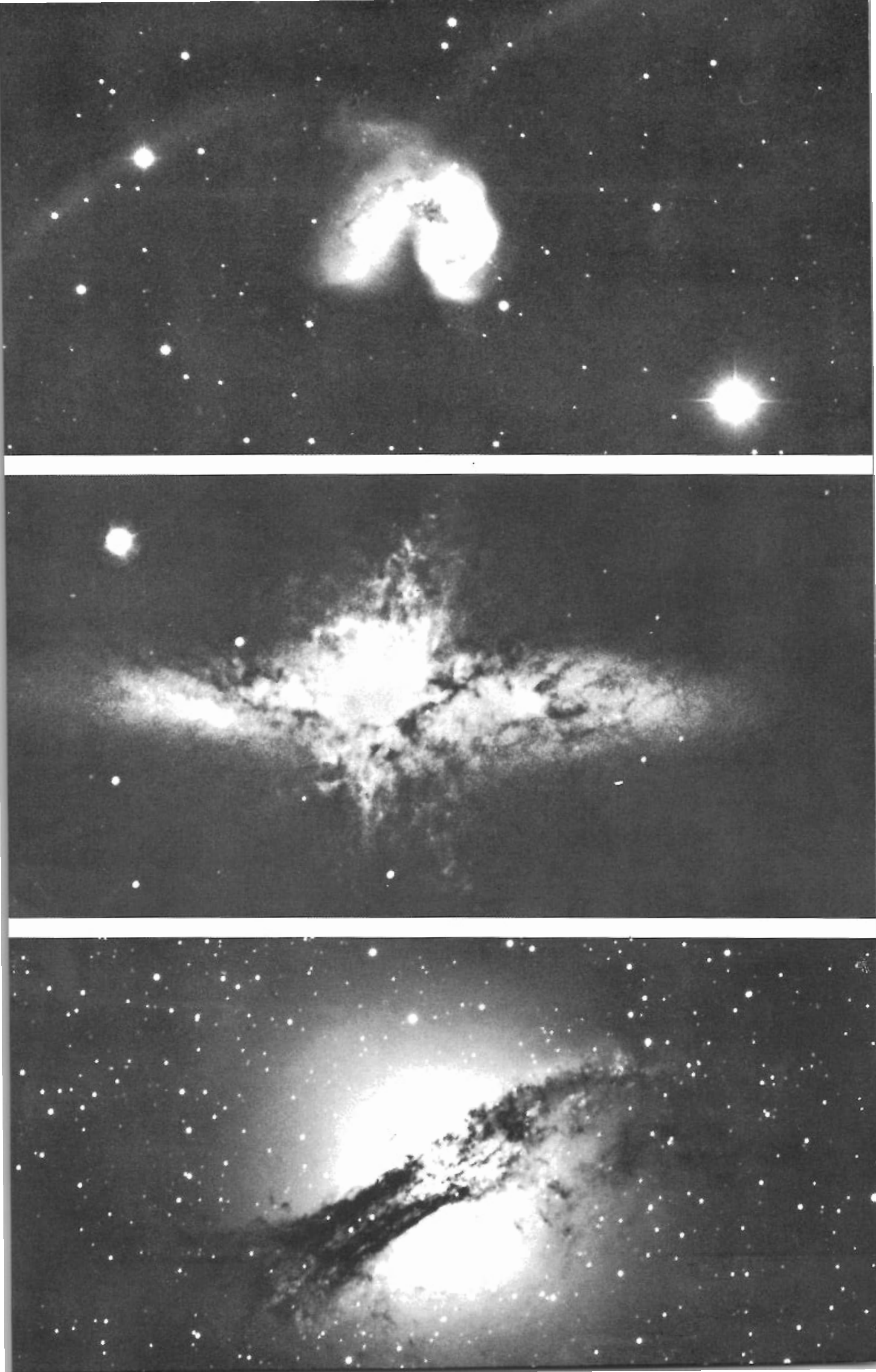






1	4
2	5
3	6

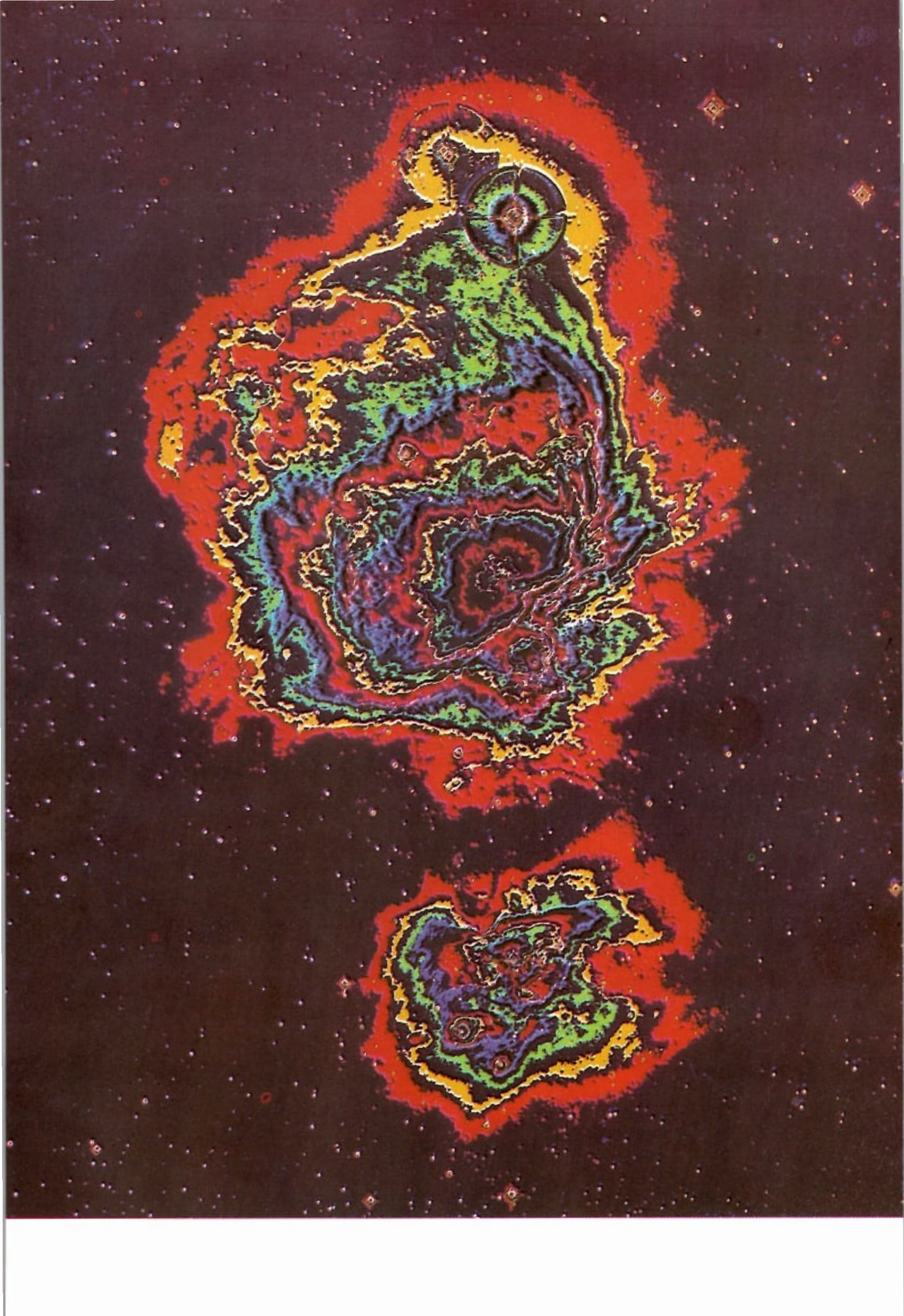
Tafeln 62 und 63 1 Komet Morehouse, 1908 III (↑ 10.2.3.13., Abb. 10.2.3-12) – **2** Spiralstruktur des Milchstraßensystems nach radioastronomischen Untersuchungen der 21-cm-Linie; Lage des galaktischen Zentrums (+) und der Sonne (⊙) (↑ 10.3.4.1.) – **3** Radiogalaxis NGC 4489 (Typ E 1), 16 Mill. Lichtjahre entfernt; der Materiestrahl aus dem Kerngebiet ist deutlich sichtbar – **4** Zusammenstoßende Galaxien NGC 4038 und 4039, ihre interstellare Materie wird dabei zu einer Quelle starker Radiostrahlung – **5** Explodierende Galaxis NGC 3034 (Typ I), 10 Mill. Lichtjahre entfernt; die ausgeschleuderten Massen streben nach oben und unten aus dem Kerngebiet – **6** Radiogalaxis NGC 5128 (Typ Ep), 12 Mill. Lichtjahre entfernt; die Breite der Absorptionsstreifen läßt auf hohe Turbulenz schließen (↑ 10.4.2.4.)





1 | 2

Tafeln 64 und 65 1 Großer Orionnebel (↑ 10.3.4.1.) –

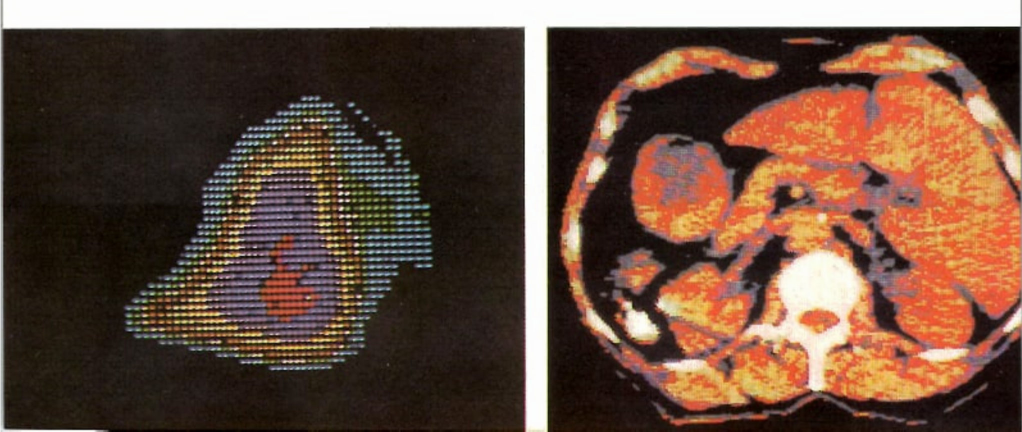


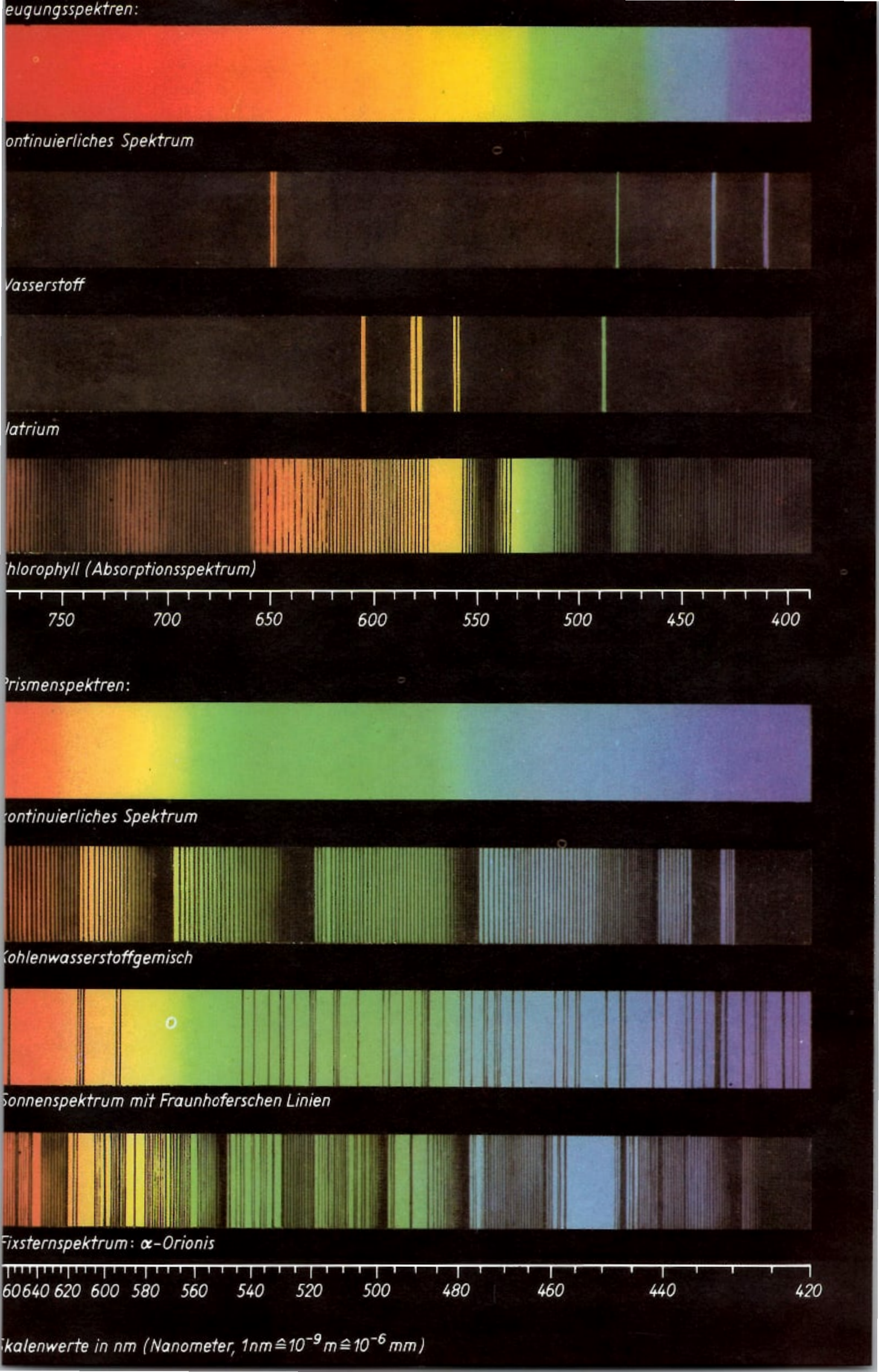
2 Darstellung einer Aufnahme des Großen Orionnebels mit Hilfe der Falschfarbentechnik

Maske 1
Farbbelichtung
Ma
Maske 2
Farbfoto-
papier
Glasplatte

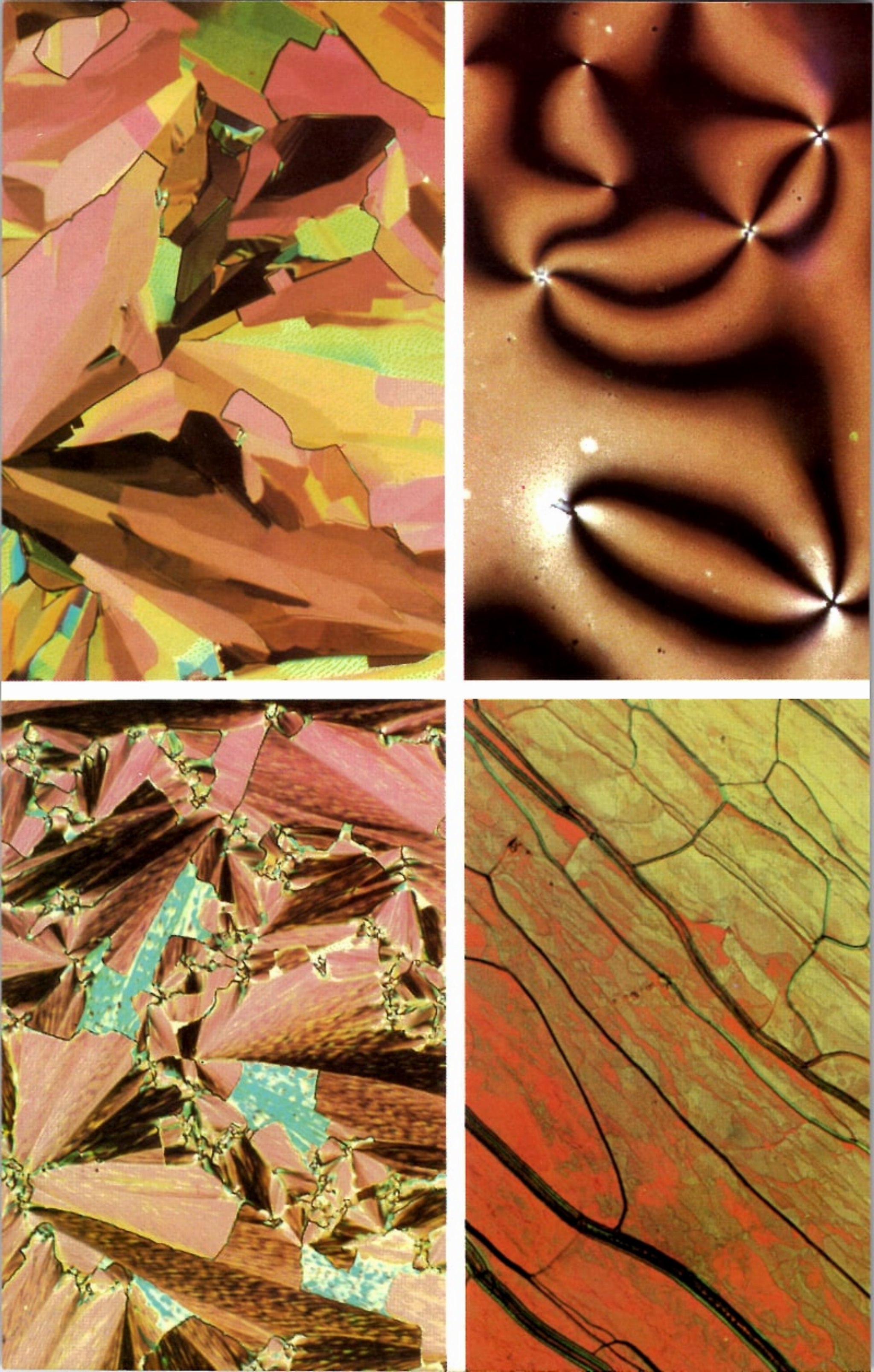
1 2 3
4 5
6 7 | 8

Tafeln 66 und 67 **1** bis **4** Elektronenmikroskopische Aufnahmen von Magnetstrukturen in einer 60 nm dicken NiFe-Schicht (Vergrößerung 5000fach); **1**, **2** und **3** entstanden bei unterschiedlicher Lage der Kontrastblende, im Farbsynthesebild **4** ist die Magnetisierungsstruktur der Domänen sehr gut erkennbar (↑ 8.6.4.4.) – **5** Schematische Darstellung der Aufnahmetechnik des Farbsynthesebilds – **6** Leberszintigramm (↑ Tafel 11) – **7** Emissionscomputertomogramm (Körperquerschnitt) des menschlichen Körpers in Höhe der Leber (↑ Tafel 11) – **8** Spektren (↑ 1.1.4.2.)



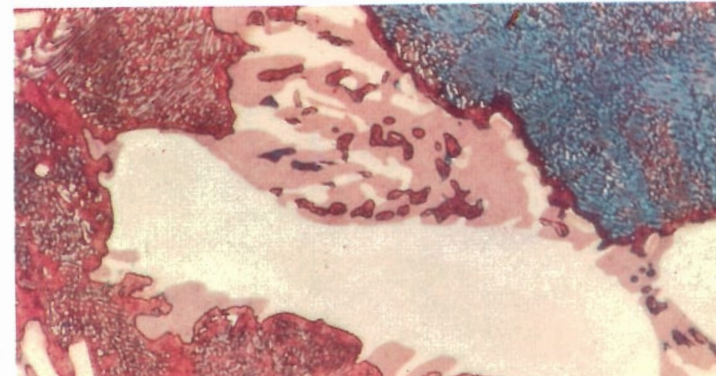


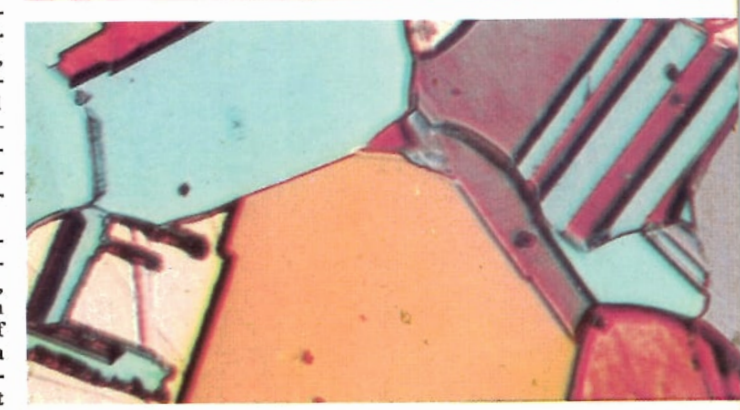
eugungsspektren:
ontinuierliches Spektrum
Vasserstoff
latrium
hlorophyll (Absorptionsspektrum)
750
700
650
600
550
500
450
400
Prismenspektren:
ontinuierliches Spektrum
Kohlenwasserstoffgemisch
Sonnenspektrum mit Fraunhoferschen Linien
Fixsternspektrum: α-Orionis
60640 620 600 580 560 540 520 500 480 460 440 420
Skalenwerte in nm (Nanometer, 1nm ≙ 10^-9 m ≙ 10^-6 mm)

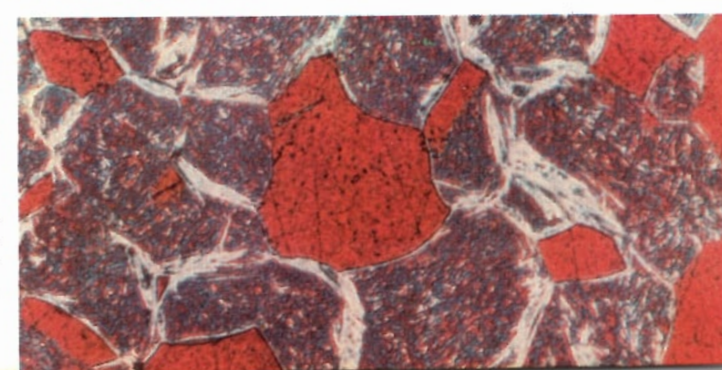


			5
1	2		6
			7
3	4		8

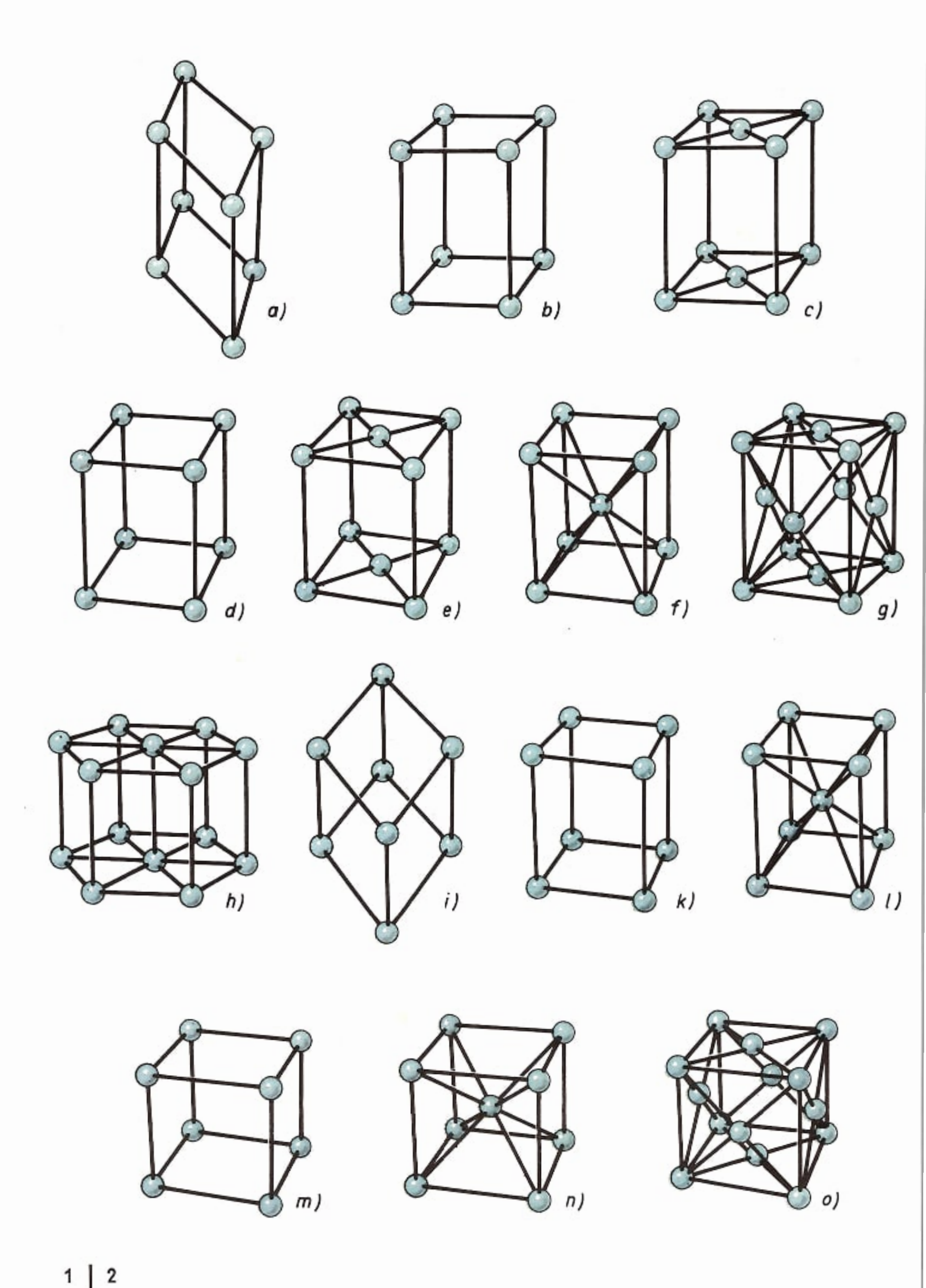




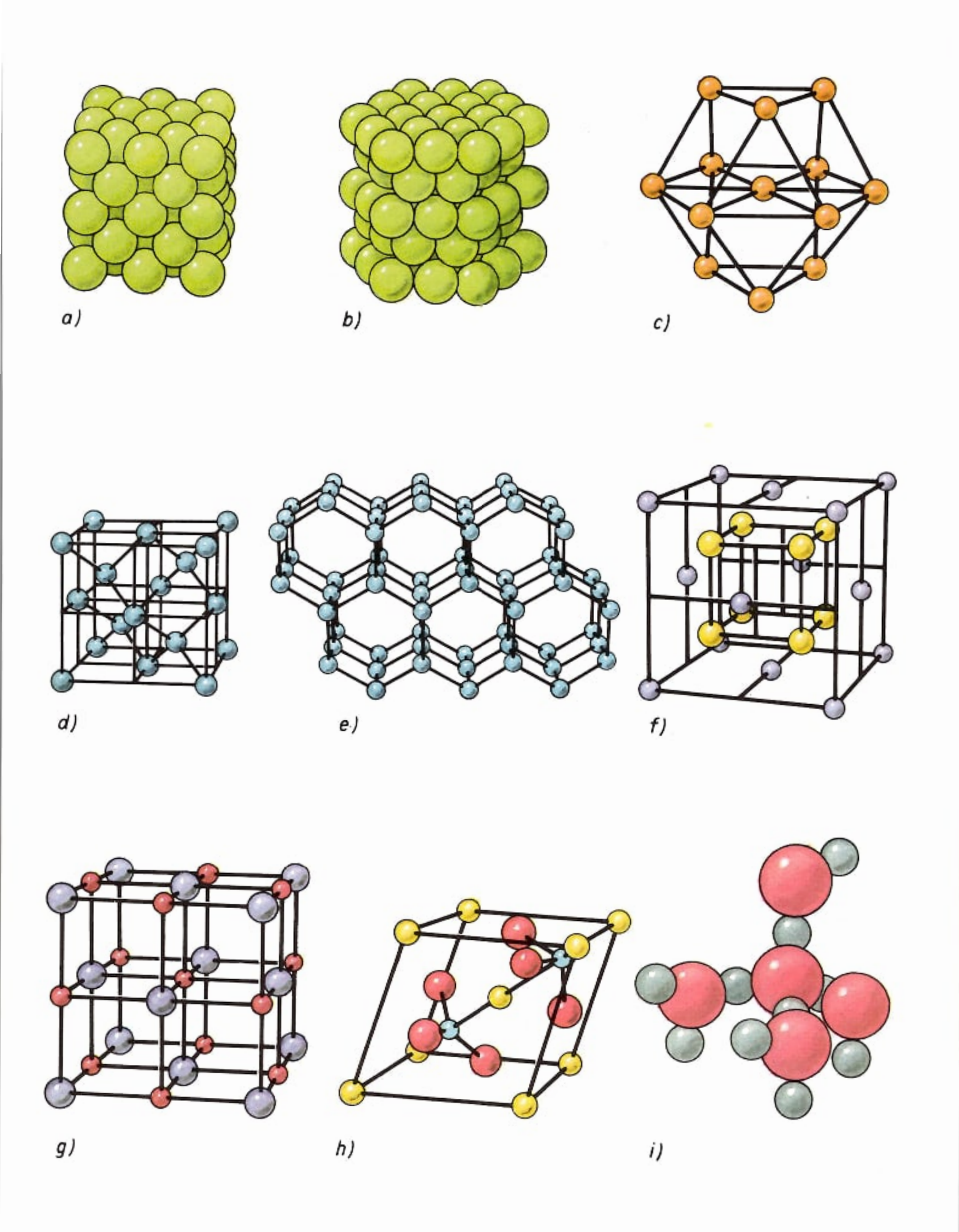




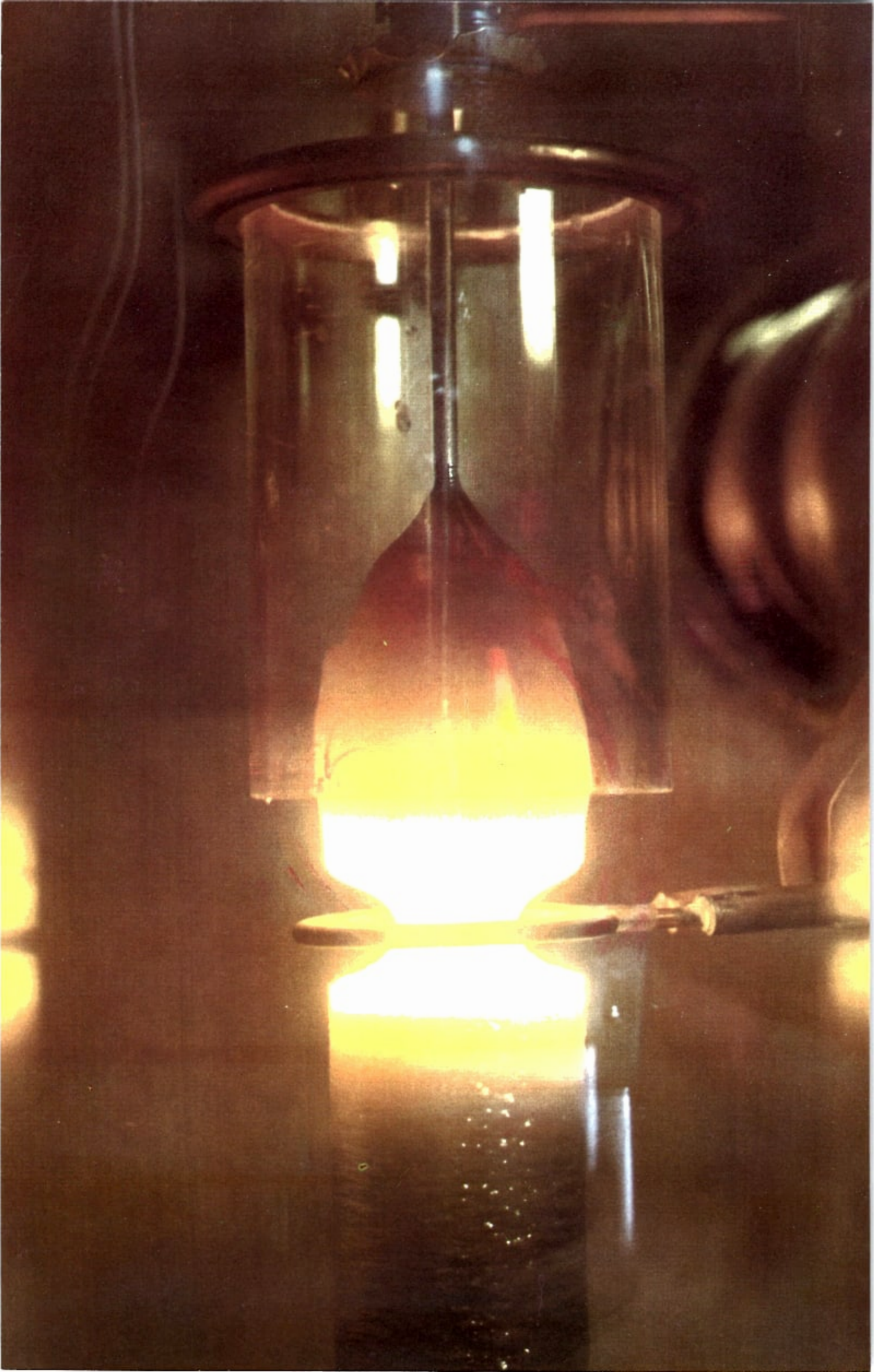
Tafeln 68 und 69 (↑ 6.3.2.1.4.)
1 Smektischer Flüssigkristall, G-Phasentyp, Mosaiktextur 3-n- Hexyl-6-[4-n-pentyloxyphenyl]-pyridazin (83 °C, Vergrößerung 50fach) – **2** Nematischer Flüssigkristall, Schlierentextur, 4-n-Butylcyclohexansäure-4-n-octyloxyphenyl-ester (71 °C, Vergrößerung 70fach) – **3** Smektischer Flüssigkristall, C-Phasentyp, durchbrochene Fächertextur, 3-n-Hexyl-6-[4-n-pentyloxyphenol]-pyridazin (90 °C, Vergrößerung 50fach) – **4** Cholesterinischer Flüssigkristall, planare Textur mit öligen Streifen, Mischung aus fünf verschiedenen Substanzen (20 °C, Vergrößerung 50fach) – **5** Schweißeisen mit Natriumthiosulfat geätzt – **6** Gußeisen 5 min lang bei 300 °C an Luft oxydiert; Eisencarbid *weiß*, Ferrit *blau* und *braun*, Eisenphosphid *graubraun* – **7** Kupfer mit Natriumthiosulfat geätzt – **8** Stahl mit 0,2 % C, 17 % Cr, 1 % Mo von 1 200°C in Öl abgeschreckt und 5min lang bei 500 °C in Luft oxydiert; *rote Grundmasse* Ferrit, *weiße Kristalle* Martensit (Vergrößerung 500fach)



Tafeln 70 und 71 1 Die 14 Translationsgitter: *a*) triklin; *b*) einfach monoklin; *c*) flächenzentriert monoklin; *d*) einfach-rhombisch; *e*) basisflächenzentriert-rhombisch; *f*) raumzentriert-rhombisch *g*) allseitig flächenzentriert-rhombisch; *h*) hexagonal; *i*) rhomboedrisch; *k*) einfach tetragonal *l*) raumzentriert-tetragonal; *m*) einfach kubisch; *n*) raumzentriert-kubisch; *o*) flächenzentriert-kubisch

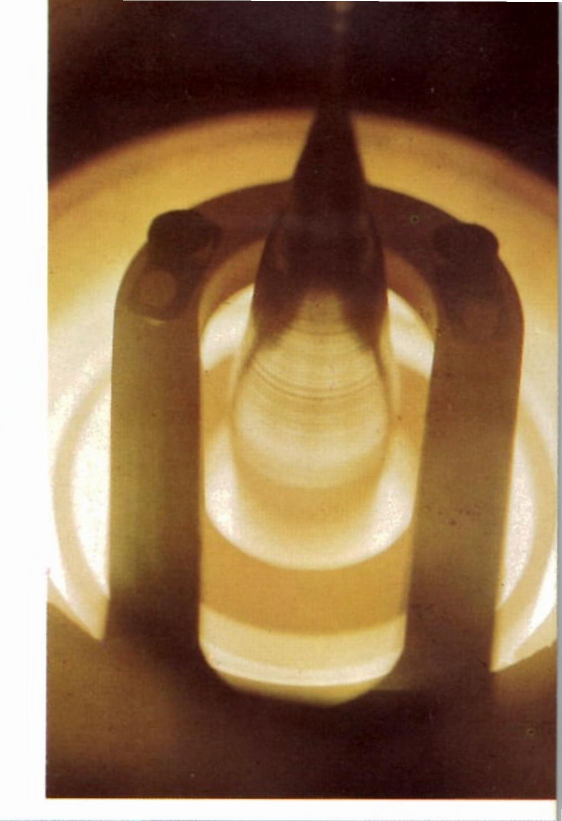


2 Modelle einiger Kristallgitter; *a*) kubisch dichteste Kugelpackung; *b*) hexagonal dichteste Kugelpackung; *c*) Magnesiumtyp (A 3); *d*) Diamanttyp (A 4); *e*) Graphittyp (A 9); *f*) Fluorittyp (C 1); *g*) Steinsalztyp (B 1); *h*) Kalzittyp (G 1); *i*) tetraedrische Nahordnungsstruktur der H_2O-Moleküle in Eis und Wasser (↑ 6.4.1.2.)



1 | 2
  | 3

Tafeln 72 und 73 (↑ 6.4.2.1.) **1** Kristallherstellung mit Hilfe des Schwebezonenverfahrens; Aufheizung eines Siliciumstabs mit einer hochfrequenzgespeisten Spule in einem schmalen Bereich – **2** Blick auf einen wachsenden Siliciumeinkristall – **3** Auswahl synthetischer Einkristalle, die hauptsächlich nach dem Čzochralskiverfahren in dem Spezialbetrieb für einkristalline Werkstoffe „Cristalec" des Forschungszentrums CNRS Grenoble hergestellt worden sind; *oben links* Halbleiterkristalle von Germanium und Silicium, *darunter* bzw. *rechts* Einkristalle verschiedener Metalle, in der *Mitte* rote Stäbe für Rubinlaser, *links unten* schön gefärbte künstliche Halbedelsteine, *unten rechts* Einkristalle magnetischer Materialien



1 | 3
2 | 4

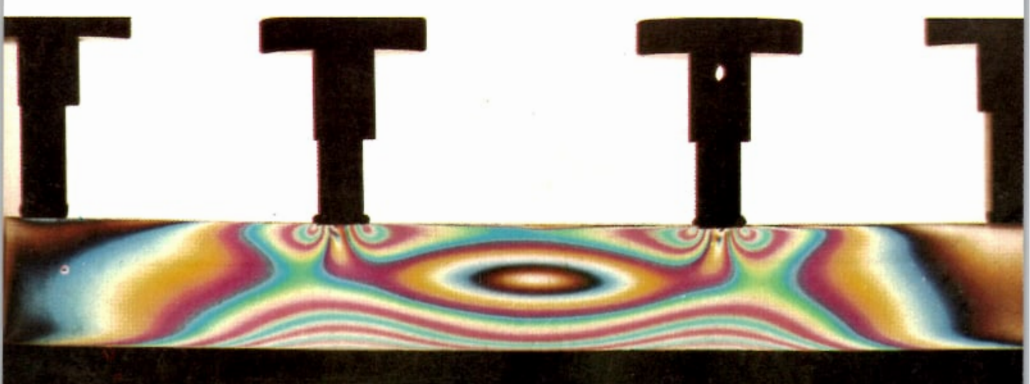
Tafeln 74 und 75 (↑ 6.4.) **1** Kristalle von gediegenem Gold auf Quarz – **2** Natürliche Silberkristalle, in blätterförmigen Skeletten gewachsen – **3** Natürlicher Barytkristall (Spaltrhomboeder) auf Fluorit (Flußspat) aufgewachsen – **4** Weiße Schwerspatkristalle zusammen mit gelben Flußspatwürfeln und dunklen Würfeln des Bleiglanzes





1	2	
3	4	9
5	6	
7	8	10

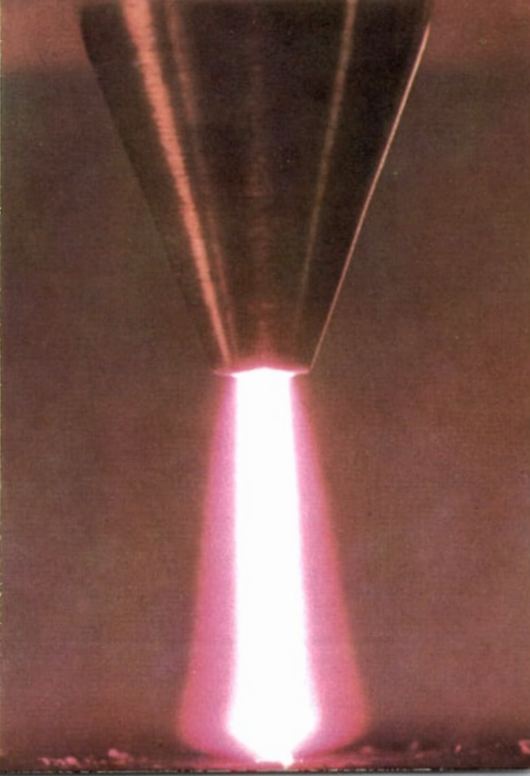
Tafeln 76 und 77 **1** Realgar – **2** Kobaltblüte – **3** Fluorit – **4** Topas – **5** gediegenes Silber – **6** Aragonit – **7** Schwerspat – **8** gediegenes Arsen – **9** Spannungsoptische Aufnahme der Zwischenwand eines Generatorgehäuses – **10** Spannungsoptische Aufnahme eines Fundamentbalkens unter Druckbelastung



1 | 4
2 3 | 5

Tafeln 78 und 79 **1** Matrixspeicherebene in Flachbauweise mit vier Drähten je Kern (↑ 8.7.5.2, Abb. 8.7.5-1) – **2** Farbstofflaser (↑ 4.1.4.6.) – **3** Plasmaschneidanlage – **4** Anomale Glimmentladung in Argon – **5** Plasmastrahl (↑ 4.1.4.7.7.)





1a	2a	3b	4b	5b	6b	7b	8			1b	2b	3a	4a	5a	6a	7a	0	Gruppe / Periode
1 H 1,0079																	2 He 4,0026	1
3 Li 6,941	4 Be 9,0122											5 B 10,81	6 C 12,011	7 N 14,0067	8 O 15,9994	9 F 18,9984	10 Ne 20,179	2
11 Na 22,9898	12 Mg 24,305											13 Al 26,9815	14 Si 28,0855	15 P 30,9738	16 S 32,06	17 Cl 35,453	18 Ar 39,948	3
19 K 39,0983	20 Ca 40,08	21 Sc 44,9559	22 Ti 47,90	23 V 50,9414	24 Cr 51,996	25 Mn 54,9380	26 Fe 55,847	27 Co 58,9332	28 Ni 58,70	29 Cu 63,546	30 Zn 65,38	31 Ga 69,72	32 Ge 72,59	33 As 74,9216	34 Se 78,96	35 Br 79,904	36 Kr 83,80	4
37 Rb 85,4678	38 Sr 87,62	39 Y 88,9059	40 Zr 91,22	41 Nb 92,9064	42 Mo 95,94	43 Tc (97)	44 Ru 101,07	45 Rh 102,9055	46 Pd 106,4	47 Ag 107,868	48 Cd 112,41	49 In 114,82	50 Sn 118,69	51 Sb 121,75	52 Te 127,60	53 I 126,9045	54 Xe 131,30	5
55 Cs 132,9054	56 Ba 137,33	57* La 138,9055	72 Hf 178,49	73 Ta 180,9479	74 W 183,85	75 Re 186,207	76 Os 190,2	77 Ir 192,22	78 Pt 195,09	79 Au 196,9665	80 Hg 200,59	81 Tl 204,37	82 Pb 207,2	83 Bi 208,9804	84 Po (209)	85 At (210)	86 Rn (222)	6
87 Fr (223)	88 Ra 226,0254	89** Ac 227,0278	104 Ku (260)	105														7

Ordnungszahl / Elementsymbol / Atomgewicht 1) — 1 H 1,0079

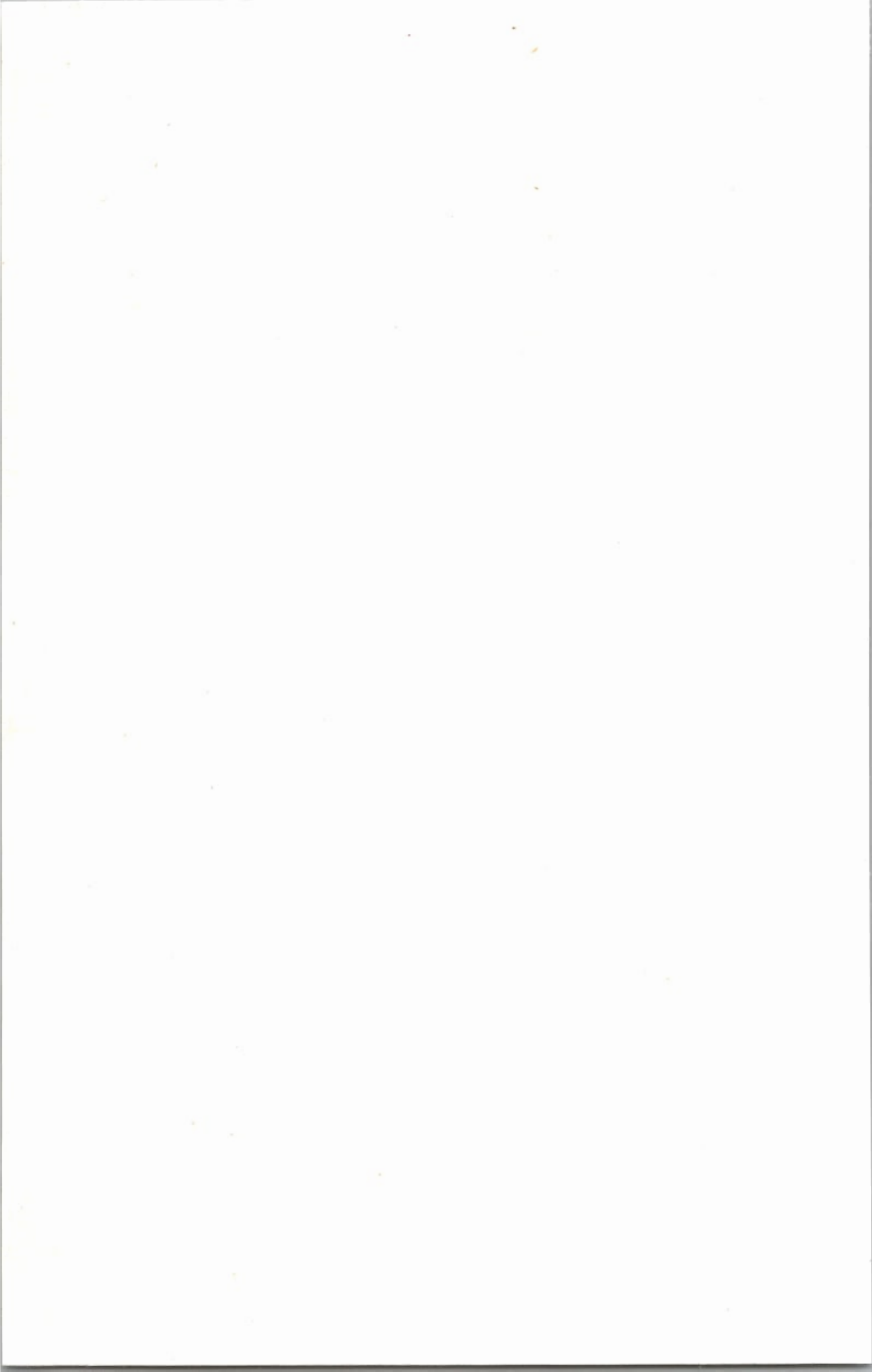
1) Zahlen in Klammern jeweils für das Isotop mit der längsten bekannten Halbwertszeit

* Lanthanoiden-elemente	58 Ce 140,12	59 Pr 140,9077	60 Nd 144,24	61 Pm (145)	62 Sm 150,4	63 Eu 151,96	64 Gd 157,25	65 Tb 158,9254	66 Dy 162,50	67 Ho 164,9304	68 Er 167,26	69 Tm 168,9342	70 Yb 173,04	71 Lu 174,97
** Actinoiden-elemente	90 Th 232,0381	91 Pa 231,0359	92 U 238,029	93 Np 237,0482	94 Pu (244)	95 Am (243)	96 Cm (247)	97 Bk (247)	98 Cf (251)	99 Es (254)	100 Fm (257)	101 Md (258)	102 No (259)	103 Lr (260)

Hauptgruppenelemente
- Oxide, säurebildend
- Oxide, basenbildend
- Oxide, amphoter
- Edelgase

Nebengruppenelemente

Tafel 80 Periodensystem der Elemente (↑ 4.2.3.4.)





Kleine Enzyklopädie

Atom- und Kernphysik